

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 10:35:17 ; Search time 2060 Seconds

(without alignments)  
11250.360 Million cell updates/sec

Title: US-09-908-988B-1

Sequence: 1 aaagagtgtagacagagtggtt.....ataaagactcaagtgctcc 1431

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 15

Total number of hits satisfying chosen parameters: 34152

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: EST:

1: em\_estbta:\*  
2: em\_estbhm:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_man:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	482	33.7	485	AK019655	AK019655 Mus muscu
2	267	18.7	461	AV006036	AV006036 AV006036
3	258	18.0	432	BB732046	BB732046 BB732046
4	168	11.7	957	AK016235	AK016235 Mus muscu
5	147	10.3	418	A2507637	A2507637 IM0349D21
6	108	7.5	316	BB140247	BB140247 BB140247

7	79	5.5	195	10	BB135648	BB135648
8	75	5.2	392	9	AV006120	AV006120
9	68	4.8	378	10	AA918753	AA918753
10	65	4.5	456	10	BE099923	BE099923
11	65	4.5	504	9	AT172700	AT172700
12	65	4.5	608	12	BE284900	BE284900
13	65	4.5	638	9	AA800245	AA800245
14	53	3.7	449	17	A2731346	A2731346
15	52	3.6	184	10	AV361069	AV361069
16	48	3.4	440	12	BG375657	BG375657
17	47	3.3	441	12	BG379943	BG379943
18	46	3.2	250	9	AV268013	AV268013
19	46	3.2	386	10	BE095943	BE095943
20	46	3.2	467	13	BM087315	BM087315
21	46	3.2	483	12	BF073506	BF073506
22	44	3.1	512	12	BF855618	BF855618
23	44	3.1	555	12	BR855618	BR855618
24	39	2.7	339	9	AA998888	AA998888
25	38	2.7	319	12	BG381122	BG381122
26	38	2.7	329	9	AT163576	AT163576
27	36	2.5	270	9	AA182748	AA182748
28	36	2.5	354	9	AA447961	AA447961
29	36	2.5	420	9	AA447961	AA447961
30	36	2.5	565	12	BF828781	BF828781
31	36	2.5	572	12	BF828778	BF828778
32	36	2.5	586	12	BF826755	BF826755
33	36	2.5	598	12	BF825744	BF825744
34	35	2.4	181	10	BB014893	BB014893
35	35	2.4	373	10	AA483766	AA483766
36	35	2.4	516	13	BM030114	BM030114
37	32	2.2	291	12	BF554230	BF554230
38	32	2.2	482	12	BF825323	BF825323
39	31	2.2	475	31	AA428229	AA428229
40	31	2.2	478	14	BQ359416	BQ359416
41	31	2.2	580	12	BF828699	BF828699
42	30	2.1	415	12	BF828153	BF828153
43	30	2.1	637	14	BM890066	BM890066
44	29	2.0	553	9	AA808889	AA808889
45	29	2.0	595	13	BJ071661	BJ071661

## ALIGNMENTS

RESULT 1	AK019655	485 bp	musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4930486E09, rRNA	linear	HTC 19-JAN-2002
LOCUS	AK019655				
DEFINITION	AK019655				
ACCESSION	AK019655				
VERSION	AK019655.1				
KEYWORDS	HTC: CAP trapper.				
SOURCE	Mus musculus (strain: C57BL/6J) adult male testis cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.				
AUTHORS	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
TITLE	Genome Res. 10 (10), 1617-1630 (2000)				
JOURNAL	20493374				
MEDLINE	11042159				
PUBMED					





/note="Vector: pSPOR1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978.

BASE COUNT 100 a 119 c 151 g 62 t

Query Match 18.0%; Score 258; DB 10; Length 432;  
Best Local Similarity 99.3%; Pred. No. 2e-112;  
Matches 408; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1017 GCCGAGATGGCTCTCTACCTCCAGCAGCAAGAGCTGATCAACAAGTCCGGGCAT 1076  
|||||  
Db 20 GCCCAGATGGCTCTCTACCTCCAGCAGCAAGAGCTGATCAACAAGTCCGGGCAT 79  
|||||  
QY 1077 GTCGAAGTGGAGTGGCAGAGAGCGCCGAGCCAGGCTATGAGACATGAGCAATTCTC 1136  
|||||  
Db 80 GTCGAAGTGGAGTGGCAGAGAGCGCCGAGCCAGGCTATGAGACATGAGCAATTCTC 139  
|||||  
QY 1137 TGTGAGCGTGGAGCAGCTGGCCGAATTTGCGAACATCGACTCCAGCGGGGGCCGC 1196  
|||||  
Db 140 TGTGAGCGTGGAGCAGCTGGCCGAATTTGCGAACATCGACTCCAGCGGGGGCCGC 199  
|||||  
QY 1197 TGGGAGTGAAGAGATGACGACATGCTTTGATGGAGAGAGGCAATGCGGGGCTGGA 1256  
|||||  
Db 200 TGGGAGTGAAGAGATGACGACATGCTTTGATGGAGAGAGGCAATGCGGGGCTGGA 259  
|||||  
QY 1257 GAGAGAGCGGCTGGAGTGGCAGAGGCTCGGCTGACACGCGGCTGATCCAGA 1316  
|||||  
Db 260 GAGAGAGCGGCTGGAGTGGCAGAGGCTCGGCTGACACGCGGCTGATCCAGA 319  
|||||  
QY 1317 GCGCAGACCCGGAAGCGGAGCCAGAGGATGCTGAGATCTCGCAGAGACACCGCGCCA 1376  
|||||  
Db 320 GCGCAGACCCGGAAGCGGAGCCAGAGGATGCTGAGATCTCGCAGAGACACCGCGCCA 379  
|||||  
QY 1377 CCAAGCTCGGCTTCCGCCCCCGGGAAGTTTCAATTAAGAGCTCAAGTG 1427  
|||||  
Db 380 CCAAGCTCGGCTTCCGCCCCCGGGAAGTTTCAATTAAGAGCTCAAGTG 430  
|||||

RESULT 4  
LOCUS AK016235/c 957 bp mRNA linear HTC 19-JAN-2002  
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:493056102; similar to RING-FINGER PROTEIN MORF, full insert sequence.  
ACCESSION AK016235  
VERSION AK016235.1 GI:12854889  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone:493056102.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

## REFERENCE

## AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagata, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, I., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojibori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staudli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weltz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S. and Hayashizaki, Y.

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Araki, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirokawa, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Konda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Schriml, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamana, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suenho-cho, Tsukuba, Ibaraki, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.resgsc.riken.go.jp, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

## COMMENT

Please visit our web site (http://genome.resgsc.riken.go.jp/) for further details.







BASE COUNT	43 a	65 c	56 g	31 t
ORIGIN				

QY 1329 AGCGGAGGCCAAGGGATGCTGAGGATCTGGCAGAAACACACGCGCCAAACACCGA 1330  
|||||  
Db 91 AGCGGAGGCCAAGGGATGCTGAGGATCTGGCAGAGACACCGCGCCACCAAGCTCGGCT 150

QY	1389	TCCCCCCCCCGGAAGTT	140
Db	151	TCCCCCCCCCGGAAGTT	169

LOCUS	DEFINITION	AV006120	AV006120	AV006120	LOCUS	DEFINITION
AV006120	392 bp	mus musculus C57BL/6J heart	mus musculus	linear	EST 25-AUG-1999	
AV006120	1020006g17, mRNA sequence.					

ACCESSION	AV0006120	
VERSION	AV006120.1	GI:4782970
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

## REFERENCE AUTHORS

TITLE	RIKEN MOUSE ESTS
JOURNAL	Unpublished (1999)
COMMENT	Contact: Chie Owa Osaka Science Laboratory

Email: genome-res@et.c.riken.go.jp  
Thermotransformation and thermoactivation of thermolabile enzymes by  
trehalose and its application for the synthesis of full length cDNA  
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))  
Transcriptional sequencing: A method for DNA sequencing using RNA  
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))  
Please visit our web site (<http://genome.etc.riken.go.jp>) for  
further details.

```
FEATURES      Location/Qualifiers
source        1..392
              /organism="Mus musculus"
              /strain="C57BL/6J"
```

[illegible]

RESULT	9
AW918753	
LOCUS	378 bp mRNA linear EST 25-MAY-2000
DEFINITION	EK350057 Rat gene index, normalized rat, norvegicus, Bello Soares Rattus norvegicus cDNA clone RGRX57 5' end, mRNA sequence.
ACCESSION	AW918753
VERSION	AW918753
KEYWORDS	AW918753.1 GI:8084538
SOURCE	EST. Norway rat.
ORGANISM	Rattus norvegicus Euteleostomi; Chordata; Craniota; Vertebrata; Euteleostomi;

REFERENCE  
1 (bases 1 to 378)  
Lee, N.H., Glodex, A., Chandra, I., Mason, T.M.,  
AUTHORS Quackenbush, J.,  
Kerlavage, A.R. and Adams, M.D.  
TITLE Rat Genome Project: Generation of a Rat EST (RESTR) Catalog & Rat  
Gene Index  
JOURNAL Unpublished (1998)  
COMMENT Contact: Lee, NH  
for Genomic Research

The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@tigr.org  
This clone is available through the ATCC, contact the ATCC  
tel#703-365-2700 for further information  
Seq primer: M3 Reverse

```

FEATURES
  source
    1. 378
    location/Qualifiers
    /organism="Rattus norvegicus"
    /db_xref="taxon:10116"
    /clone="RG1EX57"
    /clone_1fb="Rat gene index, normalized rat, norvegicus,
    Bento Soares"
    /tissue_type="mix - brain, ovary, placenta, kidney, lung,
    liver, embryo, heart, muscle, spleen"
    /lab_host="SOLR"
    /note="Vector: pluescript SK(-); Site_1: EcoRI; Site_2:
    XhoI; Estimated insert size approx.1 kb"
    87 a 100 c 122 g 69 t
BASE COUNT
ORIGIN

```

Query Match 4.8%; Score 68; DB 10; Length 378;  
 Best Local Similarity 100.0%; Pred. No. 2,1e-21;  
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 655 GCGCCCAAGAGACTGTGAGTGGCCCTGCGCCACATTATACAAAGCCAGAGAGT 714  
 |||  
 DB 58 GCGCCCAAGAGACTGTGAGTGGCCCTGCGCCACATTATACAAAGCCAGAGAGT 117  
 |||  
 QY 715 GAGCTGAG 722  
 |||  
 DB 118 GAGCTGAG 125

RESULT 10  
 BE099923/c 456 bp mRNA linear EST 13-JUN-2000  
 LOCUS UI-R-BJ1-atl-f-12-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone  
 DEFINITION  
 BE099923  
 UI-R-BJ1-atl-f-12-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone  
 BE099923  
 BE099923.1 GI:8491806  
 EST.  
 KEYWORDS  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 456)  
 REFERENCE  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 JOURNAL discovery  
 MEDLINE Genome Res. 6 (9), 791-806 (1996)  
 COMMENT 97044477

Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. The sequence tag present in the cDNA between the NotI site  
 and the oligo-dT track served to identify it as a clone from the  
 normalized AV canal at 16.5 dpc library cDNA library Preparation:  
 M.B. Soares Lab Clone distribution: clones will be available  
 through Research Genetics (www.resgen.com)  
 Seq primer: M13 Forward  
 POLYA=yes.

# FEATURES

source  
 1. 456  
 /location/Qualifiers  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-BJ1-atl-f-12-0-UI"  
 /clone\_id="UI-R-BJ1"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker. Site 1: Not I; Site 2: Eco RI; The UI-R-BJ1  
 library is a subtracted library derived from the following  
 tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV  
 canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,  
 AV canal at 15 dpc, ventricle at 13 dpc, and adult heart.  
 For a detailed description of the library from which this  
 clone was derived, please visit our web site at  
 testest.eng.uiowa.edu. The subtraction has been previously  
 described in (Bonaldo, Lennon and Soares, Genome Research  
 6:791-806, 1996)  
 TAG\_LIB=UI-R-BJ1  
 TAG\_TISSUE=AV canal at 16.5 dpc  
 TAG\_SEQ=GAACC"  
 153 c 114 g 116 t  
 73 a  
 BASE COUNT

## ORIGIN

Query Match 4.5%; Score 65; DB 10; Length 456;  
 Best Local Similarity 100.0%; Pred. No. 6,3e-20;  
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 982 AACGTGTGAGTCCGCATCCAGTCCATGAGAGAGCCGAGATGCTCTACCTCCAG 1041  
 |||  
 DB 454 AACGTGTGAGTCCGCATCCAGTCCATGAGAGAGCCGAGATGCTCTACCTCCAG 395  
 |||  
 QY 1042 CAGGC 1046  
 |||  
 DB 394 CAGGC 390

RESULT 11  
 A1712700/c 504 bp mRNA linear EST 08-JUN-1999  
 LOCUS UI-R-AF1-aau-e-02-0-UI.s1 UI-R-AF1 Rattus norvegicus cDNA clone  
 DEFINITION  
 A1712700  
 UI-R-AF1-aau-e-02-0-UI.s1 UI-R-AF1 Rattus norvegicus cDNA clone  
 A1712700  
 A1712700.1 GI:5016500  
 EST.  
 KEYWORDS  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 504)  
 REFERENCE  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 JOURNAL discovery  
 MEDLINE Genome Res. 6 (9), 791-806 (1996)  
 COMMENT 97044477

Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. The sequence tag present in the cDNA between the NotI site  
 and the oligo-dT track served to verify it as a clone from the  
 normalized AV canal at 15 dpc library cDNA library Preparation:  
 M.B. Soares Lab Clone distribution: clones will be available  
 through Research Genetics (www.resgen.com)  
 Seq primer: M13 Forward  
 POLYA=yes.

# FEATURES

source  
 1. 504  
 /location/Qualifiers  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-AF1-aau-e-02-0-UI"  
 /clone\_id="UI-R-AF1"  
 /dev stage="adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker. Site 1: Not I; Site 2: Eco RI; The UI-R-AF1  
 library is a normalized library constructed from 15 dpc  
 rat atrioventricular (AV) canal. The tag is a string of 5  
 nucleotides present between the Not I site and the  
 oligo-dT track. The library was constructed as described  
 by Bonaldo, Lennon and Soares, Genome Research 6: 791-806  
 , 1996. Tissue provided by Jim Lin, Department of Biology,  
 University of Iowa.  
 TAG\_LIB=UI-R-AF1  
 TAG\_TISSUE=AV canal at 15 dpc  
 TAG\_SEQ=GAAGC"  
 166 c 132 g 124 t  
 82 a  
 BASE COUNT

[illegible]

DEFINITION	EST19742 Normalized rat heart, Bento Soares Rattus sp. cDNA clone				
ACCESSION	RHEAM89.3' end, mRNA sequence.				
VERSION	AA800245				
KEYWORDS	AA800245.1 GI:2863200				
SOURCE	EST.				
ORGANISM	Rattus sp.				
REFERENCE	Rattus sp.				
AUTHORS	Eukaryote: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Rattus.				
TITLE	1 (bases 1 to 638)				
JOURNAL	Lee,N.H., Glodet,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerpilavage,A.R. and Adams,M.D.				
COMMENT	Rat Genome Project: Generation of a Rat EST (RESt) Catalog & Rat Gene Index				
FEATURES	Unpublished (1998)				
source	Contact: Lee, NH				
	The Institute for Genomic Research				
	9712, Medical Center Drive, Rockville, MD 20850, USA				
	Tel: (301)-838-3529				
	Fax: (301)-838-0208				
	Email: nhlee@tigr.org				
	Seq primer: M13-21.				
	Location/Qualifiers				
	1..638				
	/organism="Rattus sp."				
	/db_xref="ATCC (inhost):2007238"				
	/db_xref="taxon:10118"				
	/clone="RHEAM89"				
	/clone_lib="Normalized rat heart, Bento Soares"				
	/note="Organ: heart; Vector: pT73Pac; Site_1: EcoRI; Site_2: NotI"				
BASE COUNT	105 a	221 c	165 g	147 t	
ORIGIN					
Query Match	4.5%	Score 65;	DB 9;	Length 638;	
Best Local Similarity	100.0%;	Pred. No. 7.1e-20;			
Matches	65;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Oy	982	AAGCTGGTGAATCCGCGCATCCAGTCCATGAGAGAGCGCAGATGGCTCTACCTCCAG	1041		
Db	439	AAGCTGGTGAATCCGCGCATCCAGTCCATGAGAGAGCGCAGATGGCTCTACCTCCAG	380		
Oy	1042	CAGGC	1046		
Db	379	CAGGC	375		
RESULT 14					
AZ731346					
LOCUS					
DEFINITION	RPC1-24-117M3.TV RPC1-24 Mus musculus genomic clone RPC1-24-117M3,				
ACCESSION	AZ731346				
VERSION	AZ731346				
KEYWORDS	GSS.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryote: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.				
AUTHORS	1 (bases 1 to 449)				
TITLE	Zhao,S., Nieman,W., Malek,J., Shatsman,S., Akintre,B., Levins,M., Tsegaye,G., Geer,K., Krol,M., Shavatsbeyn,A., Gebregeorgis,E., Russell,D., de Jong,P. and Fraser,C.M.				
JOURNAL	Mouse BAC End Sequences from Library RPC1-24				
COMMENT	Unpublished (1999)				
	Other_GSSs: RPC1-24-117M3.TV				
	Contact: Shaying Zhao				
	Department of Eukaryotic Genomics				
	The Institute for Genomic Research				
	9712 Medical Center Dr., Rockville, MD 20850, USA				
	Tel: 301 838 0200				
	Fax: 301 838 0208				



Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: mscares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized rat heart pool library cDNA library preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)  
Seq primer: M13 Forward  
POLYA=Yes.

#### FEATURES

##### Source

Location/Qualifiers  
1. 440  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-CS0-btd-h-03-0-UI"  
/clone\_11b="UI-R-CS0"  
/dev\_stage="ADULT"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site.1: Not I; Site.2: Eco RI; The UI-R-CS0 library is a normalized library constructed from the following rat heart tissues: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, adult day 200. For a detailed description of the library from which this clone was derived, please visit our web site at [ratest.eng.uiowa.edu](http://ratest.eng.uiowa.edu). The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
TAG\_11b="UI-R-CS0"  
TAG\_RISUSE="rat heart pool"  
TAG\_SEQ="ATAGATATAC"  
BASE COUNT 69 a 144 c 113 g 114 t  
ORIGIN

Query Match 3.4%; Score 48; DB 12; Length 440;  
Best Local Similarity 100.0%; Pred. No. 8.7e-12;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 999 CATCCAGTCATGAGAGCGCAGATGCTCTACTCCAGCAGGC 1046  
|||||  
Db 440 CATCCAGTCATGAGAGCGCAGATGCTCTACTCCAGCAGGC 393

RESULT 17 441 bp mRNA linear EST 12-MAR-2001  
LOCUS BG379943/c  
DEFINITION UI-R-CS0-btd-h-01-0-UI.s1 UI-R-CS0 Rattus norvegicus cDNA clone  
ACCESSION BG379943  
VERSION BG379943.1 GI:13304415  
KEYWORDS EST.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
REFERENCE 1 (bases 1 to 441)  
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
COMMENT Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250  
Fax: 319 335 9565  
Email: mscares@blue.weeg.uiowa.edu  
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)  
Seq primer: M13 Forward  
POLYA=Yes.

#### FEATURES

##### Source

Location/Qualifiers  
1. 441  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-CS0-btd-h-01-0-UI"  
/clone\_11b="UI-R-CS0"  
/dev\_stage="ADULT"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site.1: Not I; Site.2: Eco RI; The UI-R-CS0 library is a normalized library constructed from the following rat heart tissues: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, adult day 200. For a detailed description of the library from which this clone was derived, please visit our web site at [ratest.eng.uiowa.edu](http://ratest.eng.uiowa.edu). The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
TAG\_SEQ="None found"  
BASE COUNT 69 a 148 c 110 g 114 t  
ORIGIN

Query Match 3.3%; Score 47; DB 12; Length 441;  
Best Local Similarity 100.0%; Pred. No. 2.6e-11;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1048 AAGGAGCTGATCACAAGTCGGGGCAATGTCGAGGTGAGCTGGC 1094  
|||||  
Db 396 AAGGAGCTGATCACAAGTCGGGGCAATGTCGAGGTGAGCTGGC 350

RESULT 18 250 bp mRNA linear EST 05-NOV-1999  
LOCUS AV268013  
DEFINITION AV268013 RIKEN full-length enriched, adult male testis (DH10B) Mus musculus cDNA clone 4930532616 3', mRNA sequence.  
ACCESSION AV268013  
VERSION AV268013.1 GI:6256050  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 250)  
AUTHORS Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Horii, F., Ishii, Y., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kojima, Y., Koya, S., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Owa, C., Ozawa, Y., Saito, H., Sano, M., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Takahashi, F., Tateo, M., Tomimaga, N., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yasunishi, A., Yokota, T., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
TITLE RIKEN Mouse ESTs (Konno, H., et al. 1999)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222

Fax: 81-45-503-9216  
 Email: genome-res@sc.riken.go.jp  
 URL: http://genome.gsc.riken.go.jp/  
 Sasaki, N., Izawa, M., Watabiki, M., Ozawa, K., Tanaka, T., Yoneda, Y., Matsura, S., Carninci, P., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
 Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)  
 Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
 Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

#### FEATURES

Source  
 Location/Qualifiers  
 1..250  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="4930532616"  
 /clone\_lib="RIKEN full-length enriched, adult male testis (DH10B)"  
 /sex="male"  
 /tissue\_type="testis"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGGAGAGAGAGATCCAGAGACCTCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGGAGAGATTCGACTTATTTAATTAATATCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plasmid pUC19(+) after bulk excision from Lambda FLX I. Cloning sites, 5' end: SalI; 3' end: BamHI."  
 BASE COUNT  
 ORIGIN  
 65 a 66 c 75 g 44 t  
 Query Match 3.2%; Score 46; DB 9; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-11;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1306 TCTGATCCAGAGCGACACCGAGCGGAGCGGATGCTGAG 1351  
 ||||||||||||||||||||||||||||||||||||||||  
 Db 127 TCTGATCCAGAGCGACACCGAGCGGAGCGGATGCTGAG 172  
 RESULT 19  
 BE095943/c 386 bp mRNA linear EST 12-JUN-2000  
 LOCUS BE095943  
 DEFINITION UT-R-BU0-apt-e-03-0-UI s1 UT-R-BU0 Rattus norvegicus cDNA clone  
 ACCESSION BE095943  
 VERSION BE095943.1 GI:8486874  
 KEYWORDS EST.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 1 (bases 1 to 386)  
 Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE  
 JOURNAL  
 MEDLINE  
 COMMENT  
 Normalization and subtraction: two approaches to facilitate gene discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 97044477  
 Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized ganglia library cDNA Library Preparation: M.B. Soares Lab Clone distribution: Clones will be available through Research Genetics (www.resgen.com)  
 Seq primer: M13 Forward  
 POLA=yes

#### FEATURES

Source  
 Location/Qualifiers  
 1..386  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UT-R-BU0-apt-e-03-0-UI"  
 /clone\_lib="UT-R-BU0"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; The UT-R-BU0 library is a subtracted library derived from a mixture of eye and ganglia tissues. For a detailed description of the library from which this clone was derived, please visit our web site at ratseq.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
 TAG\_LIB=UT-R-BU0  
 TAG\_TISSUE=ganglia  
 TAG\_SEQ=CGACAGC  
 BASE COUNT  
 ORIGIN  
 61 a 131 c 93 g 101 t  
 Query Match 3.2%; Score 46; DB 10; Length 386;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-11;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1096 GGAGCGCGGAGCGAGCGATGAGAGCATGAGCATCTCTGTGA 1141  
 ||||||||||||||||||||||||||||||||||||||||  
 Db 346 GGAGCGCGGAGCGAGCGATGAGAGCATGAGCATCTCTGTGA 301  
 RESULT 20  
 BM087315 467 bp mRNA linear EST 19-NOV-2001  
 LOCUS BM087315  
 DEFINITION 500054 MARC 2B0V Bos taurus cDNA 5', mRNA sequence.  
 ACCESSION BM087315  
 VERSION BM087315.1 GI:16997943  
 KEYWORDS EST.  
 SOURCE cow.  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.  
 1 (bases 1 to 467)  
 Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrnenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G., Pertea, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and Keeler, J.W.  
 Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle



JOURNAL	Genome Res. 11 (4), 626-630 (2001)
MEDLINE	21180013
COMMENT	Contact: Smith TPL

Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4386  
Fax: 402 762 4390  
Email: [SmithT@ars.ars.gov](mailto:SmithT@ars.ars.gov)

email: smitrn@mail.marc.usda.gov  
Single pass sequencing. Bases called and alt\_trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -mismatch 12 options.  
PCR primers

FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCACATCAGCAGC  
Plate: 137 row: H column: 12  
Seq primer: ATTTAGTGATCAGCTATAG

URES	Location/Qualifiers
source	1. .467

```

/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV Sport6; Site_1: NotI; Site_2: SalI
Library made from pooled tissue from testis, thymus,
semitendinosus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."

```

BASE COUNT	113 a	112 c	172 g	70 t
ORIGIN				

Query Match 3.2%; Score 46; DB 13; Length 467;  
Best Local Similarity 100.0%; Pred. No. 8.1e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 257 TGGAGAGCAGCTCATTTGCCCATCTGCCGTGAGATGTTCCAA 302  
 |||||  
 Db 236 TGGAGAGCAGCTCATTTGCCCATCTGCCGTGAGATGTTCCAA 281

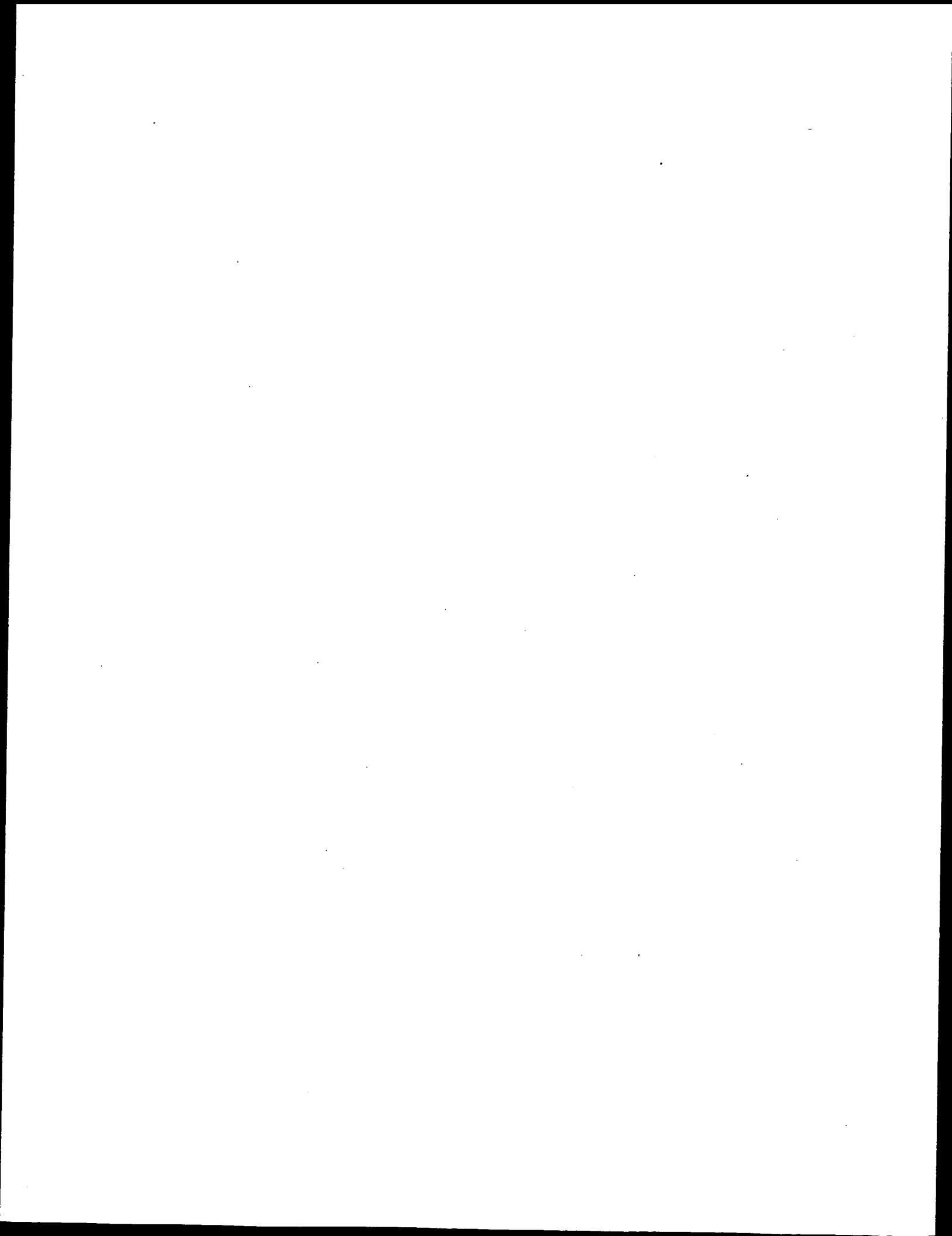
Search completed: December 3, 2002, 12:05:11  
Job time : 2081 secs



LOCUS AA800245 638 bp mRNA linear EST 30-APR-1998  
 DEFINITION EST189742 Normalized rat heart, Bento Soares Rattus sp. cDNA clone  
 RHEAM89 3' end, mRNA sequence.  
 ACCESSION AA800245  
 VERSION AA800245.1 GI:2863200  
 KEYWORDS EST.  
 SOURCE Rattus sp.  
 ORGANISM Rattus sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 638)  
 AUTHORS Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,  
 Kerlavage, A.R. and Adams, M.D.  
 TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat  
 Gene Index  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Lee, NH  
 The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-3529  
 Fax: (301)-838-0208  
 Email: nhlee@tigr.org  
 Seq primer: M13-21.  
 FEATURES  
 source Location/Qualifiers  
 1.638  
 /organism="Rattus sp."  
 /db\_xref="ATCC (inhost):2007238"  
 /db\_xref="taxon:10118"  
 /clone="RHEAM89"  
 /clone\_lib="Normalized rat heart, Bento Soares"  
 /note="Organ: heart; Vector: pT7T3Pac; Site\_1: EcoRI;  
 Site\_2: NotI"  
 BASE COUNT 105 a 221 c 165 g 147 t  
 ORIGIN

Query Match 37.6%; Score 538.2; DB 9; Length 638;  
 Best Local Similarity 91.7%; Pred. No. 2.4e-107;  
 Matches 595; Conservative 0; Mismatches 43; Indels 11; Gaps 2;

Qy 783 GGAGGAGGTGTGCCAGACCATTGAGGACAACAGCCGACAGACAGAAGCAACTGTTAAACCA 842  
 |||  
 Db 638 GGAGGAGGTGTGCCAGACCATTGAGGACAACAGCCGACAGACAGAAACAACTGTTAAACCA 579  
 Qy 843 GAGGTTTCGAGACCCCTGTGCGCGGTTTTGGAGGAGCGCAAGGGCGAACTGCTTCAAGCACT 902  
 ||  
 Db 578 GAAGTTTCGAGACCCCTGTGCGCGGTTTTGGAGGAGCGCAAGGGCGAGTTGCTTCAAGCGCT 519  
 Qy 903 GGCCCGGGAGCAGGAGGAGAAGTTGCAGCGCGTGCAGGGGCTCATCCGCCAGTACGGAGA 962  
 |||  
 Db 518 GGCTCGGGTGCAAGAGGAGAAGCTGCAGCGCGTGCAGTAGCCTCATCCGCCAGTACGGAGA 459  
 Qy 963 CCACTTGGAGGGCTCTCTAAAGCTGGTGGAGTCCGCCATCCAGTCCATGGAGGAGCCGCA 1022  
 |||  
 Db 458 CCACTTGGAGGGCTCTCTAAAGCTGGTGGAGTCCGCCATCCAGTCCATGGAGGAGCCGCA 399  
 Qy 1023 GATGGCTCTCTACCTCCAGCAGGCAAGGAGCTGATCAACAAGGTCGGGGCAATGTCGAA 1082  
 |||  
 Db 398 GATGGCTCTCTACCTCCAGCAGGCAAGGAGCTGATCAACAAGGTCGGGGCAATGTCGAA 339  
 Qy 1083 GGTGGAGCTGGCAGGACGGCCGAGCCAGGCTATGAGAGCATGGAGCAATTCTCTGTGAG 1142  
 |||  
 Db 338 GGTGGAGCTGGCTGGACGGCCGAGCCAGGCTATGAGAGCATGGAGCAATTCTCTGTGAT 279  
 Qy 1143 CGTGGAGCACGTGGCCGAAATGTTGCGAACCATCGACTTCCAGCCGGGCGCCGCTGGGGA 1202  
 |||  
 Db 278 CGTGGAGCACGTGGCCGAAATGTTGCGGACCATCGACTTCCAGCCGGGCGCCTCTGGGGA 219  
 Qy 1203 TGAAGAGGATGACGACATGGCTTTGGATGGGGAGGAGGGCAATGCGGGGCTGGAGGAGGA 1262  
 |||  
 Db 218 TGAAGAGGATGACGAGGTGACTTTGGACGGGGAAGAGGGCAACACAGGGCTGGAGGAGGA 159  
 Qy 1263 GCGGCTGGACGTGCCAGAAGGCTCAGGCCTGCACTGACCCGACTCTGATCCAGAGCGCAC 1322  
 |||  
 Db 158 GCGGCTGGACGGGCGCAGAAGGTT-----TGCACTGATCCGACTCTGATCCAGAGCGCAC 105  
 Qy 1323 ACCCGAAGCGGGAGCCAAGGGATGCTGAGGATCTGCGCAGAGACCACCGCGCCACCAAGC 1382  
 |||  
 Db 104 GCC-----CGGGAGCCAAGGGATGCTGAGGATCTGCGCAGAGACCATCGCGCTACCCATC 50  
 Qy 1383 TCGGCTTCCCGCCCCCGGAAGGTTCTCAATAAAGGACTCAAGTGTCCC 1431  
 |||  
 Db 49 TCGGCCTCCACCCCGAGGGATGGTTCTCAATAAAGAACTCATGTGTCCC 1



LOCUS AA800245 638 bp mRNA linear EST 30-APR-1998  
 DEFINITION EST189742 Normalized rat heart, Bento Soares Rattus sp. cDNA clone  
 RHEAM89 3' end, mRNA sequence.  
 ACCESSION AA800245  
 VERSION AA800245.1 GI:2863200  
 KEYWORDS EST.  
 SOURCE Rattus sp.  
 ORGANISM Rattus sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 638)  
 AUTHORS Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,  
 Kerlavage, A.R. and Adams, M.D.  
 TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat  
 Gene Index  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Lee, NH  
 The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-3529  
 Fax: (301)-838-0208  
 Email: nhlee@tigr.org  
 Seq primer: M13-21.  
 FEATURES Location/Qualifiers  
 source 1..638  
 /organism="Rattus sp."  
 /db\_xref="ATCC (inhost):2007238"  
 /db\_xref="taxon:10118"  
 /clone="RHEAM89"  
 /clone\_lib="Normalized rat heart, Bento Soares"  
 /note="Organ: heart; Vector: pT73Pac; Site\_1: EcoRI;  
 Site\_2: NotI"  
 BASE COUNT 105 a 221 c 165 g 147 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 5.93e-75 Length: 638  
 Score: 778.00 Matches: 157  
 Percent Similarity: 94.15% Conservative: 4  
 Best Local Similarity: 91.81% Mismatches: 8  
 Query Match: 40.69% Indels: 2  
 DB: 9 Gaps: 1

US-09-908-988B-2 (1-366) x AA800245 (1-638)

Qy 196 GluGluValCysGlnThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGln 215  
 Db 637 GAGGAGGTGTGCCAGACCATTGAGGACAACAGCGGCAGACAGAAACAAGTGTAAACCAG 578  
 Qy 216 ArgPheGluThrLeuCysAlaValLeuGluGluArgLysGlyGluLeuLeuGlnAlaLeu 235  
 Db 577 AAGTTCGAGACCCTGTGCGCGGTTTTGGAGGAGCGCAAGGGCGAGTTGCTTCAAGCGCTG 518  
 Qy 236 AlaArgGluGlnGluGluLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyAsp 255  
 Db 517 GCTCGGGTGCAAGAGGAGAAGCTGCAGCGCGTGCCTAGCCTCATCCGCCAGTACGGAGAC 458  
 Qy 256 HisLeuGluGlySerSerLysLeuValGluSerAlaIleGlnSerMetGluGluProGln 275  
 Db 457 CACTTGGAGGCCTCTCTAAGCTGGTGGAGTCCGCCATCCAGTCCATGGAGGAGCCGCAG 398  
 Qy 276 MetAlaLeuTyrLeuGlnGlnAlaLysGluLeuIleAsnLysValGlyAlaMetSerLys 295  
 Db 397 ATGGCTCTCTACCTCCAGCAGGCCAAGGAGCTGATCAACAAGGTCGGGGCAATGTGGAAG 338  
 Qy 296 ValGluLeuAlaGlyArgProGluProGlyTyrGluSerMetGluGlnPheSerValSer 315  
 Db 337 GTGGAGCTGGCTGGACGGCCGAGCCAGGCTATGAGAGCATGGAGCAATTCTGTGATC 278  
 Qy 316 ValGluHisValAlaGluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAsp 335  
 Db 277 GTGGAGCACGTGGCGGAAATGTTGCGGACCATCGACTTCCAGCCGGGCGCCTCTGGGGAT 218  
 Qy 336 GluGluAspAspAspMetAlaLeuAspGlyGluGluGlyAsnAlaGlyLeuGluGluGlu 355  
 Db 217 GAAGAGGATGACGAGGTGACTTTGGACGGGGAAGAGGGCAACACAGGGCTGGAGGAGGAG 158  
 Qy 356 ArgLeuAspValProGluGlySerGlyLeuHis 366  
 Db 157 CGGCTGGACGGGCCAGAA-----GGTTTGAC 131



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 3, 2002, 13:08:10 ; Search time 2056 Seconds  
(without alignments)  
2883.049 Million cell updates/sec

Title: US-09-908-988b-2  
Perfect score: 1912  
Sequence: 1 MNFTVGFKPLIGDAHNMNDL.....EGNAGLEPERLDVPEGSLH 366

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+P2n.model -DEV=xlh  
-O=/gen2.1/USPRO/spool/US09908988/runat\_26112002\_112155\_11286/app.query.fasta\_1.519  
-DB=EST -QFMT=fastlap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LIST=45  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi -ALIGN=20 -MODE=LOCAL  
-DOALIGN=200 -THR.SCORE=PCT -THR.MAX=100 -THR.MIN=0 -MAXLEN=2000000000  
-OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09908988 -ECGN\_1\_1.763 -runat\_26112002\_112155\_11286 -NCP=6 -ICPU=3  
-NO\_XLPXY -NO\_XMAP -LANG=EUCL -NEG.SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT=120  
-WARN.TIMEOUT=30 -THREX=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_iny:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vtc:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_man:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	970	50.7	595	13	Bj071661
2	900	47.1	555	12	BF855618
3	885	46.3	912	13	BF758535
4	863	45.1	516	13	BM030114
5	850	44.5	512	12	BF855619
6	849.5	44.4	868	12	BF764060
7	825.5	43.2	793	13	BF743358
8	824	43.1	1105	12	BF761218
9	813	42.5	694	10	BF852026
10	786	41.1	615	13	Bj033623
11	779.5	40.8	815	9	AL668403
12	778	40.7	638	9	AA800245
13	756	39.5	628	13	Bj059953
14	733.5	38.4	1161	12	BG153322
15	726	38.0	608	12	BF284900
16	702	36.7	495	13	BM486618
17	697	36.5	1065	17	CNS04GLR
18	695.5	36.4	1092	13	B114696
19	693	36.2	776	13	B114212
20	687	35.9	552	13	Bj072360
21	684.5	35.8	790	12	BE674823
22	683.5	35.7	582	13	Bj070440
23	683	35.7	658	14	BM726338
24	681	35.6	983	12	BE205857
25	670	35.0	579	13	Bj074770
26	668.5	35.0	1081	12	BF309583
27	629	32.9	373	10	AA483766
28	629	32.9	672	10	BB662321
29	627	32.8	578	10	AW918753
30	624	32.6	573	9	AA443443
31	617	32.3	721	12	BF765856
32	614	32.1	600	13	B1891964
33	608.5	31.8	842	17	CNS02TVC
34	605	31.4	610	17	Bj500196
35	599.5	31.4	610	17	DR161887
36	596.5	31.2	682	12	BG769996
37	595	31.1	464	9	A1644642
38	593	31.0	485	11	AK019655
39	586	30.6	795	12	BE894237
40	570	29.8	637	14	BM890066
41	557.5	29.2	688	9	AJ442216
42	547	28.6	504	9	A1712700
43	542.5	28.4	531	9	A1877673
44	536.5	28.1	508	12	BF724411
45	535.5	28.0	538	9	A1958735

## ALIGNMENTS

RESULT 1  
Bj071661  
LOCUS Bj071661 595 bp mRNA linear EST 11-DEC-2001  
DEFINITION laevis CDNA clone X1095b10 5', mRNA sequence.  
ACCESSION Bj071661  
VERSION Bj071661.1 GI:17501850  
KEYWORDS EST.  
SOURCE African clawed frog.  
ORGANISM Xenopus laevis  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae; Xenopodinae; Xenopus.  
REFERENCE 1 (bases 1 to 595)  
Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara

TITLE  
JOURNAL  
COMMENT

Expressed genes in X. laevis embryo  
Unpublished (2001)  
Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

FEATURES  
source

1. 595  
/organism="Xenopus laevis"  
/db\_xref="taxon:8355"  
/clone\_lib="X1095b10"  
/clone\_lib="NIBB Mochii normalized Xenopus tailbud  
library"  
/tissue\_type="whole embryo"  
/dev\_stage="stage 25"  
/note="Vector: pBSRN3; Site-1: NotI; Site-2: EcoRI; cDNAs  
were oligo-dT primed and directionally cloned. Staging  
according to Nieuwkoop and Faber. Library is subtracted  
and was constructed by N. Garrett and A.M. Zorn,  
(Wellcome/CRC Institute)."

BASE COUNT 145 a 158 c 162 g 130 t

## ALIGNMENT SCORES:

Pred. No.: 4,39e-96 Length: 595  
Score: 970.00 Matches: 176  
Percent Similarity: 95.85% Conservative: 9  
Best Local Similarity: 91.19% Mismatches: 8  
Query Match: 50.73% Indels: 0  
Gaps: 13

US-09-908-988b-2 (1-366) x BJO71661 (1-595)

1 MetAspNhehrValGlyPheIyProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
15 ATGAATTTTCGGTGGGTTTCACACCTTTGGCCGAGACACCCAGACCATGAGACAACTG 74  
21 GtLysGlnLeuIleCysProIleCysLeuGlnMetPheSerIyProValIleLeu 40  
75 GAGAAGCAGCTCATCTGCCCATCTGCTGAGAGATGTTCCACCAACCTGTGATCTCTG 134  
41 ProCysGlnHisAsnLeuCysArgIyCysAlaAsnAspValPheGlnAlaSerAspPro 60  
135 CCTGGCCAAACACACCTGTCGCCAAATGTGCCAATGATATATCCAGGCTCTAATCT 194  
61 LeuTrpGlnSerArgIySerThrThrValSerSerGlyArgPheArgCysProSer 80  
195 CTATGGCAGTCCCGTGGTTCACACAGTCTCTCTGGGGGCGTTTTCGTTGCCCTTCG 254  
81 CysArgHisGluValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100  
255 TCCCGCCAGAGGTGGTCTGCAGACAGACGGGGTGTACGGGCTTCAGAGGAACCTTCTG 314  
101 ValGluAsnIleIleAspIleTyrLysGlnGlnSerSerArgProIleuHisAlaIyAla 120  
315 CTGGAAAAATATCTTACATTTTACAAAGAGTCAAGGCGCCATCAAGCCCAAAATCT 374  
121 GlnGlnHisLeuMetCysGlnGlnHisGlnAspGlyLysIleAsnIleTyrCysLeuSer 140  
375 GAGCAGCAGCTCATGTGTGAAGCAGCAGGATGAGAAGATTACATCTTACGCTGAGC 434  
141 CysGluValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGluVal 160  
435 TCGGAGGTCCCTTACCTGCTCCATGTGCAAAAGTCTTTGGGGCCCAAGAGAGCTGAGGTG 494  
161 AlaProLeuProThrIleTyrLysArgGlnLysSerGlnLeuSerAspGlyIleAlaMet 180  
495 GCCCTCTTTCAGCATCTACAAAGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 554

QY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThr 193  
Db 555 CTAGTGGAGGAAATGACCGGATCCAGCAGCATCATCTACT 593

RESULT 2  
LOCUS BF855618 555 bp mRNA linear EST 16-JAN-2001  
DEFINITION RCO-FN0204-161100-031-d06 FN0204 Homo sapiens cDNA, mRNA sequence.  
VERSION BF855618 GI:12243362  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,  
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.

TITLE  
JOURNAL  
MEDLINE  
COMMENT  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RCO&t2=RCO-FN0204-  
161100-031-d06&t3=2000-11-16&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 51  
High quality sequence stop: 120.

## FEATURES

source

1. 555  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="FN0204"  
/dev\_stage="Adult"  
/note="Organ: prostate\_normal; Vector: puc18; Site-1: SmaI  
; Site-2: SmaI; A mini-library was made by cloning  
products derived from ORESPES PCR (U.S. Letters Patent  
application No. 196,716 - Ludwig Institute for Cancer  
Research) profiles into the pUC 18 vector. Reverse  
transcription of tissue mRNA and cDNA amplification were  
performed under low stringency conditions."

BASE COUNT 113 a 149 c 167 g 126 t

## ALIGNMENT SCORES:

Pred. No.: 1.9e-88 Length: 555  
Score: 900.00 Matches: 168  
Percent Similarity: 97.16% Conservative: 3  
Best Local Similarity: 95.45% Mismatches: 4  
Query Match: 47.07% Indels: 1  
Gaps: 12

US-09-908-988b-2 (1-366) x BF855618 (1-555)

QY 27 ProIleCysLeuGlnMetPheSerLysProValIleLeuProCysGlnHisAsnLeu 46  
Db 554 CCCATCTGCGTGGAGATGTCTCCAAACAGTGTGATCTGCTCCCTGCCACACAACTCG 495  
QY 47 CysArgLysCysAlaAsnAspValPheGlnAlaSerAspProLeuTrpGlnSerArgGly 66



Db 494 TGCCGCAATGTCAGACAGCTCTCCAGGCTCGATCCTATGAGCAGTCCGGG 435  
 Oy 67 SerThrTrValSerSerGlyGlyArgPheArgCysProSerCysArgHisGluVal 86  
 Db 434 TCCACACCTGTCTCTTCAGAGAGCGGTTTCCGCTCCATCGTCAGAGGCTGAGGTTGTC 375  
 Oy 87 LeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuValGluAsnIleIleAsp 106  
 Db 374 CTGGACAGACAGCTGTCTACGCGCTCGACGCAACCTGCTAGTGGAGACATTTATCGAC 315  
 Oy 107 IleTyrGlyGlnGluSerSerArgProLeuHisAlaLysAlaGluGlnHisLeuMetCys 126  
 Db 314 ATTTACAAAGCAGAGATCATCCAGCGCTGCACCTGCAAGGCTGAGCAGCAGCTCATGTGC 255  
 Oy 127 GluGluHisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThrCys 146  
 Db 254 GGGGAGCAGAAAGAGAGATCATATTACTGCTGAGTCGAGGCGGACCTGTC 195  
 Oy 147 SerLeuCysLysValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIle 166  
 Db 194 TCTCTCTGCAAGGTCTTCGCTGCCACAGAGACTGTGAGTGGCCACATGCCACCAT 135  
 Oy 167 TyrLysArgGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAsp 186  
 Db 134 TACACAGCCCAAGAGTGTGAGCTCAGCATGCGATGCGATGCTGTGCGAGGCAATGAC 75  
 Oy 187 ArgValGlnAlaValIleThrGlnMetGluGlu-ValCysGlnThr 201  
 Db 74 CCGGTGCAAGCATGATCATCAGATGAGAGCGGTGTGCCACT 29  
 RESULT 3 912 bp mRNA linear EST 25-SEP-2001  
 LOCUS B1758535 603022861F1 NIH\_MGC\_114 Homo sapiens CDNA clone IMAGE:5193326 5',  
 DEFINITION mRNA sequence.  
 ACCESSION B1758535  
 VERSION B1758535.1 GI:15750113  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1 (bases 1 to 912)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLML1483 row: j column: 15  
 High quality sequence start: 25  
 High quality sequence stop: 890.  
 Location/Qualifiers  
 1. 912  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5193326"  
 /clone\_1bp="NIH\_MGC\_114"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6  
 male brains, age range 23-27 yo. Library is oligo-dT  
 primed and directionally cloned (EcoRV site is destroyed  
 upon cloning). Average insert size 1.5 kb, insert size  
 range 1-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 019. Note:

BASE COUNT 236 a 232 c 266 g 178 t this is a NIH\_MGC Library.  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,99e-86 Length: 912  
 Score: 885.00 Matches: 166  
 Percent Similarity: 84.15% Conservative: 41  
 Best local Similarity: 67.48% Mismatches: 34  
 Query Match: 46.29% Indels: 5  
 DB: 13 Gaps: 3  
 US-09-908-988b-2 (1-366) x B1758535 (1-912)  
 Oy 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluLysGlnLeuIleCysProIleCys 29  
 Db 167 CTGATCCAGGATGGGAATCCATGGGAACCTTGGACAAACAGCTGATCTGCCCTATCTGC 226  
 Oy 30 LeuGluMetPheSerLysProValIleLeuProCysGlnHisAsnLeuCysArgLys 49  
 Db 227 CTGGAGATGTTTACCAAGCAGCTGTCATCTTGCCGTCAGCAGCAACCTGTCCGGAG 286  
 Oy 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGlnSerArgGlySerThr 69  
 Db 287 TGTGCCAATGACATCTTCCAGCTGCAAAATCCCTACTGGACACCGCGGACCTCAGTG 346  
 Oy 70 ValSerSerGlyLysArgPheArgCysProSerCysArgHisGluValAlaLeuAspArg 89  
 Db 347 TCCATGTCTGAGGCGCGTTTCCGCTGCCACACTGCCGACAGAGTATCATGATCGT 406  
 Oy 90 HisGlyValTyrGlyLeuGlnArgAsnLeuValGluAsnIleIleAspIleTyrLys 109  
 Db 407 CACGGAGTGTACGGCTCGAGAGAACCTGCTGTGAGAAATCATCGACATCTACAA 466  
 Oy 110 GlnGlu---SerSerArgProLeuHisAlaLysAlaGluGlnHisLeuMetCysGluGlu 128  
 Db 467 CAGAGTGTCTCAGTCCGCGCTG-----CACAAGGCGAGTCAACCCATGTGAAGAG 520  
 Oy 129 HisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThrCysSerLeu 148  
 Db 521 CACGAAGATGAGAAATCAACATCTACTGTCACGCTGTAAGTGGCCACCTGCTCCATG 580  
 Oy 149 CysLysValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIleTyrLys 168  
 Db 581 TCCAAAGTGTGTGGATCCACCAAGGCTCGAGGTGGCCCATTTGCAGAGTGTTCAG 640  
 Oy 169 ArgGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgVal 188  
 Db 641 GCACAAAAGACTAAGTGAATTAATCTGATCTCCATGCTGGGGGGAATGACCGTGTG 700  
 Oy 189 GlnAlaValIleThrGlnMetGluValCysGlnThrIleGluAsp-AsnSerArgAr 208  
 Db 701 CAGACCATCATCTCAGCTGAGGATTCCTCCGTCAGACCAAGGAGAAACATCTCCACA 760  
 Oy 208 GcInLysGlnLeuLeuAsnGlnArgPheGluThrLeuCysAlaValLeuGluGluArgly 228  
 Db 761 GGTAAAGGAAGACTGACCGCAAGATTGACCTTATGCTGATCCCTGGATGAGAA 820  
 Oy 228 SgLYLLeuLeuGlnAlaLeuAlaArgGluGlnGluLysLeuGlnArgValArgG1 248  
 Db 821 AAGTGAATCTCTCCAGAGATCAACGAGAGCGGAGAAAGCTTAGTATCT---CGAGG 877  
 Oy 248 YLeuIleArgGlnTyr 253  
 Db 878 CTTATCTCAGCAGTAC 893  
 RESULT 4 516 bp mRNA linear EST 05-NOV-2001  
 LOCUS BM030114 488634 MARC 2BOV Bos taurus CDNA 5', mRNA sequence.  
 DEFINITION  
 ACCESSION BM030114  
 VERSION BM030114.1 GI:16743684  
 KEYWORDS EST.

SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT

cow.  
Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.  
1 (bases 1 to 516)  
Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Kohler,G.A., Chitko-McKown,C.G., Petter,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keeler,J.W.  
Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle  
Genome Res. 11 (4), 626-630 (2001)  
21180013  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smithemail.marc.usda.gov  
Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -mismatch 12 options.  
PCR primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCAGTCACGACG  
Plate: 116 row: J column: 19  
Seq primer: ATTAGTCACACTATAC.  
Location/Qualifiers  
1. 516  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lib="MARC 2B0Y"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI; Library made from pooled tissue from testis, thymus, semitendinosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."  
BASE COUNT 119 a 137 c 171 g 89 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.95e-84 Length: 516  
Score: 863.00 Matches: 168  
Percent Similarity: 98.84% Conservative: 2  
Best Local Similarity: 97.67% Mismatches: 2  
Query Match: 45.14% Indels: 0  
DB: 13 Gaps: 0

US-09-908-988b-2 (1-366) x BM030114 (1-516)

QY 111 GluserserArpProleuHsiAlaLysAlaGluGlnHisLeuMetCysGluGlnHisGlu 130  
|||  
Db 1 GAGGCCCTCCCGCCACATGTCATCCAGGCTAGACGACCTCATGTGTGAGGAGCATGAA 60  
QY 131 AspGluLysIleasnIleTyCysLeuSerCysGluValProThrCysSerLeuGln 150  
|||  
Db 61 GATGAGAGATCATATCTACTGCTGTGAGTGGCAAGTGGCCCACTGCTCTCTGCAAG 120  
QY 151 ValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIleTyLysArgGln 170  
|||  
Db 121 GTCTTGTGTCACCAAGAGCTGTGAGTGGCCACACGTCACCATTTACAAAGCCAG 180  
QY 171 LysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGlnAspArgValGlnAla 190  
|||  
Db 181 AAGAGTAGCTCAGGAGTGCATGCGATGCTGTGCGCAATGACCGGTGTCAACCA 240  
QY 191 ValIleThrGlnMetGluGluValCysGlnThrIleGluAspAsnSerArgArgGlnLys 210  
|||  
Db 241 GTGATCACAGATGAGGAGGTGTGCCAGCATCGAGAAACAGCCGAGGAGCAGAAAG 300

QY 211 GlnLeuLeuAsnGlnArgPheGluThrLeuCysAlaValLeuGluGlnArgLysGlyGlu 230  
|||  
Db 301 CAGTTGTGAACACAGAGGTTTGAAGGCTGTGTGACGATCTGTGAGAGAGCGGAAGGTGAG 360  
QY 231 LeuLeuGlnAlaLeuAlaArgGluGlnGluLysLeuGlnArgValArgLysLeu 250  
|||  
Db 361 CTGCTGAGGCGCTGCGCCGAG 420  
QY 251 ArgGlnTyGlyAspHisLeuGlnGluGlnSerLysLeuValGlnSerAlaIleGlnSer 270  
|||  
Db 421 CGCCAGTAGGAGAGACCATCTGAGGCTCTCTTAAGCTGTGAGTGTGACATGCTCATC 480  
QY 271 MetGluGluProGlnMetAlaLeuTyLeuGlnGln 282  
|||  
Db 481 ATGAGAGAGCGCAGATGCGCTTACTGACAGCAG 516

RESULT 5  
BF855619/c  
LOCUS  
DEFINITION  
RCO-FN0204-161100-031-e06 FN0204 Homo sapiens cDNA, mRNA sequence.  
ACCESSION  
BF855619  
VERSION  
BF855619.1 GI:12243363  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 512)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
CONTACT: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Rui Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC06t2-RC0-FN0204-161100-031-e06&ts=2000-11-16&tl=1)  
Seq primer: puc 18 forward  
High quality sequence start: 42  
High quality sequence stop: 512.  
Location/Qualifiers  
1. 512  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="FN0204"  
/dev\_stage="Adult"  
/note="Organ: prostate-normal; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
BASE COUNT 104 a 138 c 157 g 113 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5.16e-83 Length: 512  
Score: 850.00 Matches: 164  
Percent Similarity: 96.51% Conservative: 2

Best Local Similarity: 95.35% Mismatches: 3  
 Query Match: 44.46% Indels: 3  
 DB: 12 Gaps: 0

US-09-908-988b-2 (1-366) x BF855619 (1-512)

QY 27 ProileCysLeuGluMetPheSerIysProValIleLeuProCysGlnHisAsnLeu 46  
 |||||  
 DB 511 CCCATCTGCTGGAGATGTTCTCAACACAGTGGTATCTGCGCTGCCAACACACCTG 452  
 QY 47 CysArgLysCysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGlnSerArgGly 66  
 |||||  
 DB 451 TGCCCAATATGTCACACACAGCTCTTCAGGCTCGAATCTCTATGGCAGTCCGGGGC 392  
 QY 67 SerThrThrValSerSerGlyGlyArgPheArgCysProSerCysArgHisGluVal 86  
 |||||  
 DB 391 TCCACACAGCTGTCTTCAGAGAGGCCCTTCCGCTGCCATCGTGCAGAGCATGAGTGTG 332  
 QY 87 LeuAspArgHisGlyValIleArgLysLeuGlnArgAsnLeuValGluAsnIleIleAsp 106  
 |||||  
 DB 331 CTGACACAGACAGCTGTCTACGGCTGCAGACCAACCTCTAGTGGAGAACATATGAC 272  
 QY 107 IleTyrLysGlnGluSerSerArgProLeuHisAlaLysGluGlnHisIleuMetCys 126  
 |||||  
 DB 271 ATTTCACACAGAGCTATCCAGCGCTGCTCAAGCTGAGCAGCACTCATGTG 212  
 QY 127 GlnGluHisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThrCys 146  
 |||||  
 DB 211 GAGGACATGAGAGAGAGATCATATTTACTGCTGAGCTGCGAGGTGCCACCTG 152  
 QY 147 SerLeuCysLysValPheGlnAlaHisLysAspCysGluValAlaProLeuProThrIle 166  
 |||||  
 DB 151 TCTCTCTGCAAGGCTCTCGGCTGCCACAGAGCTGAGGTGGCCCACTGCCACCAT 92  
 QY 167 TyrLysArg-GlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGlyAsn 186  
 |||||  
 DB 91 TACTAACGCGCCCGAAGAGAGAGCTCAGCGATGCGATGCGATGCTGTGGCAGCAATG 32  
 QY 186 sPArGyAlGlnAla-ValIleThrGlnMet 195  
 |||||  
 DB 31 ACCGCTGCAAGCATGTGATCACACCATG 2  
 RESULT 6  
 BG764060 868 bp mRNA linear EST 15-MAY-2001  
 LOCUS 602737069F1 NIH\_MGC\_49 Homo sapiens cDNA clone IMAGE:4862321 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG764060  
 VERSION BG764060.1 GI:14074713  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 868)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: ATCC/DCMP/DMP  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 plate: LICM1722 row: j column: 18  
 high quality sequence stop: 863.  
 Location/Qualifiers  
 1. 868  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

/clone="IMAGE:4862321"  
 /clone\_id="NIH\_MGC\_49"  
 /tissue\_type="melanotic melanoma, high MDR (cell line)"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site: 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCAAGAG(c). Size-selected >500bp for average insert size  
 1.8kb. Library constructed by Ling Hong in the Laboratory  
 of Gerald M. Rubin (University of California, Berkeley)  
 using 2AP-cDNA synthesis kit (Stratagene) and Superscript  
 II RT (Life Technologies). Note: this is a NIH\_MGC  
 Library."

BASE COUNT 224 a 222 c 252 g 170 t

## ALIGNMENT SCORES:

Pred. No.: 1,38e-82 Length: 868  
 Score: 849.50 Matches: 163  
 Percent Similarity: 82.58% Conservative: 41  
 Best Local Similarity: 65.99% Mismatches: 37  
 Query Match: 44.43% Indels: 6  
 DB: 12 Gaps: 2

US-09-908-988b-2 (1-366) x BG764060 (1-868)

QY 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGlnLysGlnLeuIleCysProIleCys 29  
 |||||  
 DB 138 CTGATCCAGAGATGGAGATCCCATGGAGACTTGGAGAGAGCTGTATCTGCTATCTGC 197  
 QY 30 LeuGluMetPheSerLysProValIleLeuProCysGlnHisAsnLeuCysArgLys 49  
 |||||  
 DB 198 CTGGAGATGTTTACCAAGCAGAGCTATCTTCCCTGCGCAGACAACTGTGCCGAGAG 257  
 QY 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGlnSerArgLysIleThr 69  
 |||||  
 DB 258 TCTGCCAATGACATCTTCCAGGCTGCAATCCCTACTGAGCAGCGCGGAGCTGCTG 317  
 QY 70 ValSerSerGlyValArgPheArgCysProSerCysArgHisGluValIleLeuAspArg 89  
 |||||  
 DB 318 TCCAGTCTGTGAGGCGGTTTCCCTGCCCTGCCACCTGCGCAGAGGTGATCATGAGCT 377  
 QY 90 HisGlyValTyrGlyLeuGlnArgAsnLeuValGluAsnIleIleAspIleTyrLys 109  
 |||||  
 DB 378 CACGAGATGTAGGCTGCGAGAGAACCTGCTGTGAGACATCATCATCTACAAA 437  
 QY 110 GlnGlu---SerSerArgProLeuHisAlaLysAlaGlnGlnHisLeuMetCysGluGln 128  
 |||||  
 DB 438 CAGGATGCTCCAGTCCGCGCGCTG-----CAGAAAGGAGTCAACCCCATGTGCAAGAG 491  
 QY 129 HisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValAlaProThrCysSerIle 148  
 |||||  
 DB 492 CACGAAGATGAGAAATCAAGATCTACTCTGCTGAGAGTGGCCACCTGCTCAT 551  
 QY 148 ucysLysValPheGlnAlaHisLysAspCysGluValAlaProLeuProThrIleTyrLys 168  
 |||||  
 DB 552 GTCCAGAGGTGTTGGATCCACAAAGGCTCGAGGTGGCCCATTTGAGAGTGTCCA 611  
 QY 168 sArgGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgVal 188  
 |||||  
 DB 612 GGGACAAAGACTGACTGATTAATCTATCTCATCTGCTGCGGGGAGATGACCGGT 671  
 QY 188 IGlAlaValIleThrGlnMetGluGluValCysGlnThrIleGlnAspAsnSerArgArg 208  
 |||||  
 DB 672 GCGACCATATATCATGAGTGGAGATTCCTCGTGGAGTACCAAGAGAACATGTCACCA 731  
 QY 208 gClnLysGlnLeuLeuAsnGln-ArgPheGlnThrLeuCys-AlaValLeuGluGluArg 227  
 |||||  
 DB 732 GGTAAAGAGAGAGCTGAGCCAGACAGTTTGACAGCTTAGTGCATCTGATGAGAG 791  
 QY 228 LysGlyGlnLeuLeuGlnAlaLeuAlaArgGlnGlnGlnLysLeuGlnArgValArg 247  
 |||||  
 DB 792 AAAAGTGTGTTGGTGCAGGATCACGCGAGGAGAGACAAAGCTTAAGTTATCCGAG 851

OY	248	GLYLELLEARGLN	252	
Db	852	GGCCTCATCCAGCAG	866	
RESULT 7				
LOCUS	B1754358			
DEFINITION	B1754358	793 bp	mRNA	EST 25-SEP-2001
	603026906P1 NIH_MGC_114	Homo sapiens	cDNA clone IMAGE:5197560 5'	
ACCESSION	B1754358			
VERSION	B1754358.1	GI:15745936		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1 (bases 1 to 793)			
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: csapbs-femail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: LLAM11494 Row: k Column: 01 High quality sequence stop: 793. Location/Qualifiers 1..793 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5197560" /clone_lib="NIH_MGC_114" /lab_host="DH10B" /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 primate brains, age range 23-27 yo. Library is oligo-dt primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC library."			
BASE COUNT	201 a	208 c	228 g	156 t
ORIGIN				
Alignment Scores:				
Pred. No.:	5,14e-80	Length:	793	
Score:	825.50	Matches:	153	
Percent Similarity:	85.45%	Conservative:	35	
Best Local Similarity:	69.55%	Mismatches:	28	
Query Match:	43.17%	Indels:	4	
DB:	13	Gaps:	2	
US-09-908-988B-2 (1-366) x B1754358 (1-793)				
OY	10	LeuDeuglYaSPAlaHiAsnMeFasPaSnuLeuGlInuLysGlnLeuIlleCyProIleCyS	29	
			::	
Db	140	CTGATTCACAGATTGGGAATCCCAATCGAACAACCTTGAGAAGAACGCTGATCTGCCCTATATCTGC	199	
OY	30	LeuGlUmetPheserLySProValValIleLeuProCYsGlnHisAsnLeuCYsArgLYs	49	
Db	200	CTGGAGATGTTTACCAGCACAGTGTCTATCTTGCCGTCAGCACACACTGTGCGGAG	259	
OY	50	CysAlaSnAspValIPheGlnAlaSerAsProLeuTrpGlnSerArAglySerThrThr	69	
Db	260	TGTGCCATGACATCTTCAGGCTGCAGAAATCCCTACTCTGGACACACCGGGGCAAGCTCAGTG	319	

OY	70	ValSerSerg1ylylrPheAqgCysProSerCyArghisgluValvAlleaAsparg	89
Db	320	TTCATGCTCGAGAGCCGGTTTCCCCTGCCACTGCCACAGGTGATCATGTGATCG	379
OY	90	HISGLVAllyTrglyLeuglnAr-gAsnLeuValgluaSnlleleAspIleTylys	109
Db	380	CACGGAGGTACGGCCTGCAGAGAACCTGCTGTGGAGAACATCATCGACATCTCAAA	439
OY	110	GInGUl---SerSerArpProLeuHISAlalySaIagUGlnHisLeuetyCysGUlgu	128
Db	440	CAGAGATGCTCCAGTGCGCCCGCTG-----CAGAAGGCAGTCACCCCATGTCCAAGAG	493
OY	129	HISGLUAPGluLyIsleasnIleTyrcysLeuSerCysGUlvalPProThrCysSerleu	148
Db	494	CACGAAGATGAGAAATCAACAATCTACTGCTCTCACSTGTAGAGTGCACACTGCTCCAG	553
OY	149	CysLysValPhelgylalharstlysAspycSUlvaAlaAlaProleuprothrIleTyrys	168
Db	554	TGCMAAGTITTTGGGATCCACAAAGCCCTCGAGGTGGCCCAATGAGAGTGTCTTCCAG	613
OY	169	ArgGlnLysSerGluDeuSeraspGlyIleAlaMetLeuValaIaglysaAsparGval	188
Db	614	GGACAAAAGACTGAACTGAAATACTGATTCOCATGCTGTGGCGGGAATGACCTGTG	673
OY	189	GlAlaValIlleThrglmetgluGUlvalCysGlnThrIlegluAspaSnerArg-ar	208
Db	674	CAGACCATTATCACTCAGCTGAGGAGTTCCGCTCGAGTAGCACCAGAGAGAACGTCCACAG	733
OY	208	GlnLysGlnLeuLeuAsnGlnArGPhegluThrLeuCysAlaValleGlnGUlarG	227
Db	734	GTACACGGAGAGAGCTGAGCCAGACAGTCTGACAGCTTTGATGCCATCTTGATGAGAGAG	791
RESUT 8			
LOCUS	BG761218	1105 bp	mRNA linear EST 15-MAY-2001
DEFINITION	602718564P1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4858622 5'		
ACCESSION	BG761218		
VERSION	BG761218.1	GI:14071871	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	NIH-MGC http://mgs.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC/DCTP/DRP CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LECM1712 row: P column: 15 High quality sequence stop: 787. location/Qualifiers 1..1105 /organism="Homo sapiens" /db_xref="taxon:.9606" /clone="IMAGE:4858622" /clone_id="NIH_MGC_49" /tissue_type="melanotic melanoma, high MDR (cell line)" /lab_host="DH10B (phage-resistant)" /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(5). Size-selected >500bp for average insert size 1.8kb. library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley)		

using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library. 1"

BASE COUNT 302 a 279 c 325 g 199 t

## ALIGNMENT SCORES:

Alignment Scores: 1.29e-79 Length: 1105  
Pred. No.: 824.00 Matches: 178  
Score: 77.08% Conservative: 44  
Percent Similarity: 61.81% Mismatches: 56  
Best Local Similarity: 43.10% Indels: 11  
Query Match: 12 Gaps: 3  
DB:

US-09-908-988b-2 (1-366) x BG761218 (1-1105)

10 LeuLeuGLYASpLahIsasmetaspasleuGLYlueuileuilecysProilecys 29  
137 CTGATCCAGATGGGAATCCCATGGAGAACTTGAGAGACGCTGATGCTGCTATCTGC 196  
30 LeuGLuMetPheSerLysProValValIleuProCysGln-HisAsnLeuCySArGly 49  
197 CTGAGATGTTTACCAAGCCAGTGGTCTATCTTGGCCAGTCACAACTGTGCGGAA 256  
49 scysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGlnSerArGlySerThrTh 69  
257 GTGTGCCAATGACATCTTCCAGGCTGCAATCTCTACTGGACACCGGGCAGCTCAGT 316  
69 ValSerSerGlyLysPheArGlyCysProSerCysArGHisGluValValLeuAspAr 89  
317 GTCCAGTGTGGAGGCGCTTCCGCTCCACCTGCCG-CACGAGTGATCATGATGATCG 375  
89 GHisGlyValTyrGlyLeuGlnArGAsnLeuValGluAsnIleLeuAspIleTyrLys 109  
376 TCACGGAGTGTAGCGCTGAGAGAACTCTGTGGAGAACATCATCGACATCTTACA 435  
109 sgInGLu---SerSerArGProLeuHisAlaLysAlaGluGlnHisLeuMetCysGluG 128  
436 ACAGAGATGCTCCAGTGGCGCTG-----CAGAGGGCAGTACCCCATGTGCAAGA 489  
128 uHisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThrCysSerLe 148  
490 GCACGAGATGAGAAATCAACATCTACTGTCACGTGAGTGGCCACCTGCTCCAT 549  
148 uCysLysValPheGlyLahIsLysAspCysGluValAlaProLeuProThrIleTyrLys 168  
550 GTGCAAGTGTGGGATCCACAAAGGCTGCGAGGTGCGCCCATTCGACAGTGTCTCA 609  
168 sArgGluLysSerGluLeuSerAsp-GlyIleAlaMetLeuValAlaGluAsnAspArgV 188  
610 GGGACAAAGATGAGTGAATGACTGATCTCATGCTGTGCGGGGAAATGACCGTG 669  
188 alGlnAlaValIleThrGlnMetGluGluValCysGlnThrIleGluAspAsnSerArgA 208  
670 TGCAGACCATATCATCTAGCTGAGGAGTCTCCGTCGAGTGAACAAAGGAAACAGTCA 729  
208 rGlnLysIleLeuLeuLeuAsnGlnArGpHeGluThrLeuCysAlaValLeuGluGlu-Arg 227  
730 AGGTAAGAGAGAGAGTGTAGCCAGAAAGTTGACACGTTGATGATCATCTGATGAGAGA 789  
228 LysGlyGluLeuLeuGlnAlaLeuAlaArGluGlnGluGlu-LysLeuGlnArGValAr 247  
790 CAACGAGTGTGGTGCAGAGATCAACGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 849  
247 gGlyLeuLeuLeuGlnTyrGlyLysPheHisLeuGluGly---Ser-SerLysLeuValGlu 266  
850 AGGCTCATCCACAGAGATACCCAGAGAGGAGTGGGACAAAGTCAACAAAGCTGTGGA 909  
266 eAlaIleGln-SerMetGluGluProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGlu 285  
910 CTGCATCATCAAGTCCCTGTGGAGAGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 969  
286 LeuIleAsnLys 289

Db 970 TCATCCACAAAG 981

RESULT 9  
BB520266

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BB520266 694 bp mRNA linear EST 25-OCT-2001  
BB520266 RIKEN full-length enriched, 16 days neonate heart Mus  
musculus cDNA clone D830041C10 3', mRNA sequence.  
BB520266  
BB520266.2 GI:16443394  
EST.  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 694)  
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hangaki, T., Hara, A.,  
Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Kono, H., Kouda,  
M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Sasaki,  
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,  
D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,  
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,  
Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
Unpublished (2001)  
On Jul 28, 2000 this sequence version replaced gi:9571724.  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suenho-cho, Tsukuba, Ibaraki, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsr.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,  
M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
Wagl, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watanabe, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura,  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,  
Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamakawa, I., Aizawa,  
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
Hayashizaki, Y.  
Computational Analysis of full-length mouse cDNAs compared with  
Human Genome Sequences Mamm. Genome 12, 673-677 (2001)  
Please visit our web site (http://genome.gsc.riken.go.jp/) for  
further details.  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Location/Qualifiers  
1..694  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone\_lib="D830041C10"  
/clone\_lib="RIKEN full-length enriched, 16 days neonate  
heart"  
/tissue\_type="heart"  
/dev\_stage="16 days neonate"

## FEATURES

source

/lab\_host"DH10B"  
/note="Site\_1: SalI; Site\_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genetic Sciences Center and genome science laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGGAGAGACGGCCGCCAACACGACTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGGAGAGAGATTCGTGAATTATTAATAATATCCTCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pluescript KS(+) after bulk excision from lambda FLC I." 178 C 179 g 153 t 1 others

Alignment Scores:	
Pred. No.:	9,766-79
Score:	813.00
Percent Similarity:	89.84%
Best Local Similarity:	79.68%
Query Match:	42.52%
DB:	10
	Gaps: 2
	Length: 604
	Matches: 149
	Conservative: 19
	Mismatches: 17
	Indels: 2
	Gaps: 2

US-09-908-988B-2 (1-366) x BB520266 (1-694)

	LOCUS	DEFINITION	EST
QY	13	ASpA1Hh1s1n1e1s1p1a1n1e1u1g1u1y1g11p1e1u1l1e1c1y1p1r1o1l1e1c1y1s1e1u1g1m1e1t	32
Dp	140	GAGCAGCAGACCCATGGATACCTTGAAAGAACCACTGATCTGTCCATCTGCTTACAGATG	199
QY	33	P1e1s1e1r1y1p1r1o1v1a1l11e1l1e1u1P1r1o1C1y1s1g1n1h1s1A1s1n1e1u1c1y1s1a1r1g1y1c1y1s1A1a1s1n	52
Dp	200	TTCACCAAGCCCTGGTGCATCTTCCTTGGCACACCAACCTGGCAGAAATGGCCAGT	259
QY	53	ASp1a1P1h1e1g1n1A1a1S1e1r1s1n1P1r1o1l1e1u1T1r1p1G1n1S1e1r1A1r1g1y1S1e1r1T1h1r1V1a1S1e1r1	72
Dp	260	GACATTTTCCAGGCCCTCTAACCCTGACTTACCACCAAGAGAGAGCACACCCTGGCATCA	319
QY	73	G1Y1G1A1Y1A1r1P1h1e1a1r1g1c1y1P1r1o1S1e1r1c1y1S1A1r1g1h1S1g1u1A1Y1A1l1e1u1S1P1A1r1g1h1S1g1U1A1	92
Dp	320	GGGGGGCCGCTTCGGCTGCTCCCTCTCAACATGAGTGGTGTTAACAGACATGGGGTC	379
QY	93	T1r1G1Y1e1u1G1n1A1r1g1s1n1e1u1A1G1u1A1s1n11e1l1e1A1s1P1l1e1r1y1S1g1n1U1s1e1r	112
Dp	380	TATGGACTCAGAGGAAACCTGCTCGTGGAAACATATTGATATTCTACACAGCAGGAATCC	439
QY	113	S1e1r1A1r1P1r1o1l1e1H1s1A1a1Y1s1A1a1g1u1g1n1h1S1l1e1u1e1t1C1y1S1g1u1n1h1S1g1u1A1S1P1g1u1	132
Dp	440	ACCAAGGCCA---GAAAAAAATTGGACCCAGGCC--ATGTGTGAACAGCATGAAGAGGAA	493
QY	133	L1y1S1l1e1A1s1n11e1T1e1r1C1y1S1e1u1S1e1r1c1y1S1g1U1A1P1r1o1T1h1C1y1S1e1r1e1u1c1y1S1y1S1A1P1h1e	152
Dp	494	CGCATCAACATGATTGTTCTGAACGTGAAGGCCACCTGTTCCTTGTGCAAGGTTTTT	553
QY	153	G1Y1A1H1S1Y1S1A1S1P1C1y1S1g1U1A1A1P1r1o1l1e1u1P1r1o1T1h11e1T1r1Y1S1A1r1g1n1U1S1e1r	172
Dp	554	GGCGGCCATTAAGACTGGCAGCGTGGCTCCCGACTCAAGTGTTCACAGAGCAGAAAGTCA	613
QY	173	G1u1e1u1S1e1r1A1S1P1Y1l1e1A1a1e1t1e1u1A1A1a1g1A1y1A1s1A1S1A1r1g1A1e1A1n1A1Y1l1e	192
Dp	614	GAGCTCAGATGATGTATGTCTGTACTTGTGGACACCAAGATAGATCCAGGGTGTGATC	673
QY	193	T1h1r1G1n1e1t1G1u1g1U1A1C1y1S	199
Dp	674	AGCCAGCTGGAGACACCTGT	694
RESULT	10		
LOCUS	Bj033623	615 bp	mrna
LOCUS	Bj033623		linear
DEFINITION	Bj033623	NIBB Moch1 normalized Xenopus neurula library	Xenopus

ACCESSION	B0033623
VERSION	B0033623.1
KEYWORDS	GI:17414076
SOURCE	EST
ORGANISM	African clawed frog, <i>Xenopus laevis</i>

SOURCE	AFRICAN clawed frog
ORGANISM	Xenopus laevis

**KEYWORDS** EST

VERSION BJ03

ACCESSION BJ03

laev

REFERENCE  
1 (bases 1 to 615)  
Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and Kohara

Expressed genes in *X. laevis* embryos

COMMENT Contact: Tadasu Shin-i

Center for Genetic Resource Information  
National Institutes of Health

1111 Yata, Mishima, Shizuoka 411-8540 Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

**FEATURES**

Location/Qualifiers

source 1. .615

```
/organism="Xenopus laevis"
```

```

/wb_acl= caou:0333
/clone="XT022f1A"

```

```
/clone_lib="NIBB mochil no
```

Library

```
/cissae_type= whole embryo
/dev stage="stage 15"
```

```
/note="Vector: PBSRN3; Site
```

were oligo-dT primed and d

and was constructed by Nieuwkoop and

(Wellcome/CRC Institute).

ORIGIN	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	

ORIGIN

Alignment Scores:	
Pred. No.:	7,39e-76
Score:	786.00
Percent Similarity:	87.05%
Best Local Similarity:	75.13%
Query Match:	41,118
DB:	13
Length:	611
Matches:	140
Conservative:	23
Mismatches:	23
Indels:	2
Gaps:	2

03-03-908-988B-2 (1-366) X BU033623 (1-615)

[illegible]

	Db	398	GACCAACA---ANGTGTGATGAGCATGAAGATGAGAATAATATTATTATCTGTGCAC	454
	Qy	141	CysgluValProthrCysSerLeucCysLysValPheGlyAlaHisLysAspCysGluVal	160
	Dd	455	TGTGAAATTCCTACTCCCTCCATCGCAAAATGTTTGAGCTCACAAAATGTGAGGTT	514
	Qy	161	AlaProLeuPcOTHrIleTyrlYsArgGlnJnsSergIuleuSerAspGlyIeAlaMet	180
	Dd	515	GCCTCTCTTACACAAGTTTTCCAAGACAAGAGTCGTGACACTCTCAGATGGCATTCATA	574
	Qy	181	LeuValAlaGlyAsnAspArgValGlnAlaValIleThr	193
	Dd	575	CTGGTGGATGTAATGACAGAAATTCAGCAGCACTAGTAAC	613
	RESULT	11		
	LOCUS	AL668403		
	DEFINITION	AL668403 directional larval cDNA library Clona intestinalis cDNA		
	ACCESSION	clone 034ZAI1 5', mRNA sequence.		
	VERSION	AL668403		
	KEYWORDS	AL668403.1 GI:18129994		
	SOURCE	EST.		
	ORGANISM	Clona intestinalis.		
	AUTHORS	Clona intestinalis		
	TITLE	Eukaryote; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;		
	JOURNAL	Pneubobranchia; Clonidae; Clona.		
	COMMENT	1 (bases 1 to 815)		
	REFERENCE	Genoscope.		
	AUTHORS	Clona intestinalis directional larval cDNA library		
	TITLE	Unpublished (2002)		
	JOURNAL	Contact: Genoscope		
	COMMENT	Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seqef@genoscope.cns.fr, Web : www.genoscope.cns.fr IMPORTANT: this sequence may contain errors. The Clona intestinalis library from which the clone was isolated may be contaminated with cdnas from bacteria or other Eukarya. directional larval cDNA library originate from Dr.M.Branno, Stazione A.Dohrn, Naples, Italy, and was prepared in Pbluescript2SK+.		
	FEATURES	Location/Qualifiers		
	SOURCE	1..815		
		/organism="Clona intestinalis"		
		/db_xref="taxon:7719"		
		/clone="034ZAI1"		
		/clone_lib="directional larval cDNA library"		
		/note="Vector: pBluescript2SK+"		
	BASE COUNT	259 a 146 c 194 g 213 t	3 others	
	ORIGIN			
	Alignment Scores:			
	Pred. No.:	6.04e-75	Length:	815
	Score:	779.50	Matches:	146
	Percent Similarity:	76.25%	Conservative:	53
	Best Local Similarity:	55.94%	Mismatches:	11
	Query Match:	40.77%	Indels:	11
	DB:	9	Gaps:	4
	US-09-908-988B-2 (1-366) x AL668403 (1-815)			
	Qy	10	LeuLeuGlYasPaLa-----HISAmMeTaAspaSnDeuGlulYsgInLeuIleCysPro	27
	Dd	60	TTGATGGGNGGGCGGTGGGCCGAATCCGAGAAMVTGAAACGGCTTCGATGTGCA	119
	Qy	28	IleCysLeuGlunetPheSerLysProValValIleLeuProCysGlnHisAsnLeuCs	47
	Dd	120	ATATGCTCTTGAATCTACACCACMACCCGTTGTTATCTCCCTCCACGACATPACTTGTC	179
	Qy	48	ArgLysCysLaaAnaSPaValPheGlnAlaSerAsnProLeuTrpGlnSerArgGlySer	67
	Dd	180	AGGAATAGCGCCACGATGCTTT-----CAGAACAGGGGA---	215

Oy	68	ThThValSerSerGlyGlyArgPheArgCysProSecYsrRghISgluValLeu	87
Db	216	ACTCCATGGGGTGGCGGGAGATTCTAGGTGCCTACGGTGCTTAGCAAGTTGTRATTG	275
Oy	88	AsparghISgluValTyrgClgLeugInArqAsnleuDeuValGlusnIleIeaspIle	107
Db	276	GACCAGCATGAAGCTCTACGGCTTCACAAGAACCCTACTTGTTGAGATAATTAATGACATG	335
Oy	108	TyrLysGlnclUserSerArgProLeuHisAlalysAlaIglnGlnHISlemetCyglu	127
Db	336	TATCAACACAGAAAGGAGAAAACCAACTTTC--AAAACCTCAGAGGTTGTATATGCGAC	392
Oy	128	GIuhISgluAspgluYlysIleasnIIeTyrcYSLeuSerCYgluValProThrCYseSer	147
Db	393	CACATGTAGAAGAGAAAAGTGAATATTACTAGCTGCTGACTGTGTACGAAMCAACAGCTGTCT	452
Oy	148	LeucYlsValPheglyAlaHisLysAspcYsgluValAlaIAProLeuProThrIleTyr	167
Db	453	CTGTGCAAAAGTATTGGGGACCAACACAGCGTGTATATGCCCTAATGAGAGAAAATATAC	512
Oy	168	LysArgGlnYlsSerGluLeuSerAspGlyIleAlamethleValAlaIagLYAsnAspArg	187
Db	513	AAGAACAAAAAGAGGAATATCTGACTCTTAGCACCTCCTAGTGGCAGAAAGATCGT	572
Oy	188	ValGlnAlaValIleThrGlmetylGlnlValuAlcYsglnThrIlegLuAspAsnSerArg	207
Db	573	CTTCAGCAATATTATTTCACAAACAGAAAGAGCTGGAAAACGCGCAATGTAAACGGTCTG	632
Oy	208	ArgGlnLysGlnLeuLeuAsnGlnArqPheglunThrLeuCYsAlaValLeuGlnGluArg	227
Db	633	AAAGGAAGAAGTATTTATCGCTCGCTGGTTTGAACAACTTTCACGCCGTGTGGAGCAACGT	692
Oy	228	LysGlyGluLeuLeuGlnAlaLeuAlaArqGlnGlnGlnGlnGlnLysLeuGlnArqValArq	247
Db	693	AAGACTGARATATATGAATCAACATTAACCTTGATGTCTGAAGAAAGAACATCGATTATANA	752
Oy	248	GlyLeuIleArGlnTyrgLyAsphISleuGlnGlySerSerLyLeuValGlnSerIa	267
Db	753	GANAGTGTAAATCTTACGCTCTCAACTGTAGTCAGCAAGCAAGTATTAATGAGACCTTCA	812
Oy	268	Ile 268	
Db	813	CTT 815	
RESULT 12	AA800245/c	638 bp	mRNA linear EST 30-APR-1998
LOCUS	AA800245		
DEFINITION	EST189742 Normalized rat heart, Bento Soares Rattus sp., cDNA clone.		
ACCESSION	AA800245		
VERSION	AA800245.1	GI:2863200	
KEYWORDS	EST.		
SOURCE	Rattus sp.		
ORGANISM	Rattus sp. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Ratus.		
REFERENCE	1 (bases 1 to 638)		
AUTHORS	Lee,N.H., Glodok,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D.		
TITLE	Rat Genome Project: generation of a Rat EST (RESt) Catalog & Rat Gene Index		
JOURNAL	Unpublished (1998)		
COMMENT	Contact: Lee, NH The Institute for Genomic Research 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: (301)-838-3529 Fax: (301)-838-0208 Email: nhlee@igr.org Seq primer: M13-21. Location/Qualifiers 1..638 /organism="Rattus sp."		
FEATURES	Source		





```

RESULT 14
LOCUS      BG163322                1161 bp    mRNA    linear    EST 06-FEB-2001
DEFINITION 602338256p1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4446564 5',
ACCESSION  BG163322
VERSION    BG163322.1  GI:12670025
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
            (bases 1 to 1161)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs.r@email.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LINL at:
            http://image.llnl.gov
            Plate: LLAM10225  row: 0  column: 13
            High quality sequence stop: 639.
FEATURES
    source
        1..1161
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4446564"
            /clone_1ib="NIH_MGC_89"
            /tissue_type="hypernephroma, cell line"
            /lab_host="DH10b (phage-resistant)"
            /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally; oligo-gt primed.
            Average insert size 1.3 kb. Library enriched for
            full-length clones and constructed by Life Technologies.
            Note: this is a NIH_MGC Library."
BASE COUNT  375 a      288 c      291 g      207 t
ORIGIN
Alignment Scores:
    Score: 121e-69      Length: 1161
    Percent Similarity: 73.50      Matches: 156
    Best Local Similarity: 74.30%      Conservative: 29
    Query Match: 38.36%      Mismatches: 51
                                Indels: 14
                                Gaps: 4
US-09-908-988b-2 (1-366) x BG163322 (1-1161)
QY      1  MetAsnpheThValGlypHelysProleuLeuGlyAspAlaHisAsnMetAspAsnLeu 20
        |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::
Db      132  ATGAGCCCATCTCGAATTTACAATCTTTTCCAAAGACGACAGCATGGATTAACCTTA 191
QY      21  GluLysGlnLeuIleCysProIleCysLeuGlnMetPheSerLysProValIleLeu 40
        |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::
Db      192  GAGAACGACATCATCTGCCATCTGCTTAGAGAGTTCACGAAACCTGTGTGATTCAC 251
QY      41  ProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPhe-----GlnAl 57
        |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::
Db      252  CCTGTGAGCACAACCTGTGTAGGAATGTGCCATGATATTTCACGACAGAACAGCC 311
QY      57  aserAspProleuTrpGlnserArgLysSerThrThrValSerSerGlyLysPheAr 77
        |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::
Db      312  CTTAAACCCGTAATTTGCCACAAAGAGAGATCACCAATGATCATCAGGGCGGATTCG 371
QY      77  gCysProSerCysArgHisGlnValValLeuAspArgHisGlnValTyrGlyLeuGlnAr 97
        |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::
Db      372  CTGCCCATCTGTAGACATGAAGTGTGTTGGATAGACATGGGGTATATGACTTCAGAG 431

```

```

QY      97  gasnLeuLeuValGlnAsnIleIleaspIleTyrLysGlnGlnSerArgProleuH 117
        |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::
Db      432  GAACCTGCTGTGGAAAATATCATTTACATACAGCAGAGGTCCACCAAGCCA----- 486
QY      117  sAlaLysAlaGlnGlnHisLeuMetCysGlnGlnHisGlnAspGlnLysIleAsnIleTy 137
        |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::
Db      487  -GAAAAGAAATCCGACCAAGCCCATCTGGCAGACATGAAAGAGAGCCGATCAACATCTA 545
QY      137  rCysLeuSerCysGlnValProThrCysSerLeuCysLysValPheGlyAlaHisLysAs 157
        |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::
Db      546  CTGCTGAACAGTCGGAAGATCACCCATCTGCTGTGCAAGGGTGTGGACACACAAGA 605
QY      157  pCysGlnValAlaProleuProThrIleTyrLysArgGlnLysSerGlnLeuSerAspG 177
        |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::
Db      606  CTGCCAGGGGCTCCCTCCTCATTGATGTTCACAGACAGAAAGTGCATCATATG 665
QY      177  yLeaLamLeuValAlaGlyAsnAspArgValGln---AlaValIleThrGlnMetG 196
        |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::
Db      666  GCTCCATC-CTCGTGGGCACAAACATCAGATCCAGGGAGACAAAGCACACTGTGTA 724
QY      196  uGlnValCysGlnThrIleGlnAsp-AsnSerArgArgGlnLysGlnLeuLeuAsnGlnA 216
        |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::
Db      725  AGAACCTCTCAACTATGAGGAATGTTCCGATACACACAAAGAGCTTGTCAGAA 784
QY      216  rgPheGlnThrLeuCysAlaValLeuGlnGlnArg-----LysGlyGlnL 231
        |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::
Db      785  ACTGCTCACATCTGTTGGAAGTATAGGGGAAAGGACATGATGACACAGGATTTCA 844
QY      231  euLeuGlnAlaLeuAlaArgGln 238
        |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::  |||:::  :::::
Db      845  TTCAGGAACACTAGTAGAGAAA 867
RESULT 15
LOCUS      BF284900/c              608 bp    mRNA    linear    EST 28-NOV-2000
DEFINITION BF284900 Rat Gene Index, normalized rat, Rattus norvegicus cDNA
ACCESSION  BF284900
VERSION    BF284900.1  GI:11215970
KEYWORDS   EST.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 608)
REFERENCE  1
            Malek,R.L., Cho,J., Lee,Y., Karamycheva,S., Parvizzi,B., Pertea,G.,
            Sultana,R., Tsai,J., White,J., Quackenbush,J. and Lee,N.H.
            Generation of ESTs from Normalized Rat Embryo, Bento Soares
            Unpublished (2000)
            Other ESTs: EST350057
            Contact: Lee, NH
            The Institute for Genomic Research
            9712, Medical Center Drive, Rockville, MD 20850, USA
            Tel: (301)-838-3529
            Fax: (301)-838-0208
            Email: nhlee@tigr.org
            This clone is available through the ATCC, contact the ATCC
            tel#703-365-2700 for further information.
FEATURES
    source
        1..608
            /organism="Rattus norvegicus"
            /db_xref="taxon:10116"
            /clone="RG1EX57"
            /clone_1ib="Rat Gene Index, normalized rat, Rattus
            norvegicus cDNA"
            /tissue_type="mixed tissue"
            /lab_host="DH5-alpha"
            /note="Vector: pT37Pac; Site_1: EcoRI, Site_2: NotI;
            Combination of RSV, RBR, RRI, RLI, RPL, RLU, RRM, RRU, RSP
            , RHE, RPC, RPN"
BASE COUNT  101 a      207 c      162 g      138 t

```



	108B04 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION	AL289800
VERSION	AL289800.1 GI:8028380
KEYWORDS	GSS; genome survey sequence.
SOURCE	Tetraodon nigroviridis.
ORGANISM	Tetraodon nigroviridis Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE	1 (bases 1 to 1065) Roest-Crollius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J. Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
AUTHORS	Unpublished
TITLE	2 (bases 1 to 1065) Roest-Crollius,H., Jalllon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J. Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 1065)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-APR-2000)
COMMENT	This sequence is a single read and was generated as part of a large scale clone-and sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <a href="http://www.genoscope.cns.fr/Tetraodon">http://www.genoscope.cns.fr/Tetraodon</a> .
FEATURES	Location/Qualifiers
source	1..1065 /organism="Tetraodon nigroviridis" /db_xref="taxon:99883" /clone="108B04" /clone_lib="G"
BASE COUNT	/note="Genoscope sequence ID : C0BG108D0A02LP1-end : T7"
ORIGIN	253 a      300 c      312 g      196 t      4 others
Alignment Scores:	
Pred. No.:	1,07e-65      Length:      1065
Score:	697.00      Matches:      133
Percent Similarity:	72.77%      Conservative:      38
Best Local Similarity:	56.60%      Mismatches:      48
Query Match:	36.45%      Indels:      17
DB:	Gaps:      2
US-09-908-988B-2 (1-366) x CNS04GLR (1-1065)	
Oy	17 MetAspaSnLeungLuJySGInLeuIlleCysProIleCysteIeuclunMetPheSerLySPro 36      :::     :::     :::     :::     :::     :::     :::     :::
Db	350 ATGAGACAGCTTGAGAGACAGCTGAGCTGCCATCGTCATGTGACATGGTCACCAAAACCG 409      :::     :::     :::     :::     :::     :::     :::     :::
Oy	37 ValVallleuPrcYsgJnHisAsnLeuCysArgPyScysAlaAsnAPValpHeGln 56    :::     :::     :::     :::     :::     :::     :::     :::
Db	410 GTGGTCATCCCTGCCTCCACACACACTCTGCCGCGAGTTGTGCCACGACCCTTAGAC 469    :::     :::     :::     :::     :::     :::     :::     :::
Oy	57 AlaSerAnPrLeuTrpGlnSerArgGlySerThrValSerSerGlyGlyArgPhe 76 :::         -----CGCTCTCTGGGGGGCGCTTC 505
Oy	77 ArgcYpSeoSerCysAlaGhisgluValValleuAsparghisgluValTYrGlyLeuGln 96      :::     :::     :::     :::     :::     :::     :::     :::
Db	506 CGCTGCTCTACTCTCCGCTTTGAAGTGTCTTTACCGCACGGGTGTGCACGGGGCTGCAG 565      :::     :::     :::     :::     :::     :::     :::     :::
Oy	97 ArgAsnLeuValGluaSnlllelEaspIleTyrlYsgJngInGlnUserSerArgProLeu 116      :::     :::     :::     :::     :::     :::     :::     :::
Db	566 CGCAACCTGTTGGTGGAAAATATTATCATCATCTTCACACAAGCAGCAAGAAGCTCAGCGNC 625

OY	117	HISALALysAlagluGlnHisLeu-----MetCysgluLu	128
Db	626	AGCGGGAMACCCCAAGCAAGCTGAAGACGAAAGAGGCCAATGACCATTGTGCATATA	685
OY	129	HieGlunspgluLysIleasnIleTyrcyLeuSerCysgluValProThrCysSerLeu	148
Db	686	CACGAGGGAGAGAAGATCAACTCTACTGTGTGACTCTTCAGAGTCCCACCTGCCTACTC	745
OY	149	CysLySValPheglYAlaHisLySaSPCySGluValAlaProLeuProthrIleTyrlYS	168
Db	746	TGCAAGGCTCTTTGGCCAGCACAAAGCTGGAGAGTGCCACTTTACCGGCTCTTAACGAG	805
OY	169	ArgInLysSerGIueuSeraspGlyIleAlaMetLeuValAlaGlyAsnspaRyaVl	188
Db	806	ACCCACACAGGCGGASGTGAGCAACGCTATCGAACCTGGTGTCCAGCAATGAGCGCTTC	865
OY	189	GlnAlaValIleThrglnMetgluGluValCysGlnThrIleGluAspAsnserArGaY	208
Db	866	CAGCGCTGCTCAACCAACGATGGAAGAGCGCTGCCGCGCGGTATGAGAAAAGCTCGAAT	925
OY	209	GlnIysGlnLeuAsngInArGyphegInThrLeuCysAlaValIeugluGluArGlys	228
Db	926	GCTATGACCAAGGCTCTCCGAGAGCTTGAGCTGTGTATCTCTCTCTCGAAGR-CGCAAG	984
OY	229	GlygluLeuLeuGlnAlaLeuAlaArggluGlnGluLysLeu	243
Db	985	GCCTTCCTGCTGGAGAGATGACGAAAGAGCAGAGMCMCAAGAGTG	1029
RESULT 18	Bll114696	1092 bp	mRNA EST 26-JUN-2001
LOCUS	Bll114696	602861723F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:502097 5'	
DEFINITION		mRNA sequence.	
ACCESSION	Bll114696		
VERSION	Bll114696.1	GI:14565597	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cs9abs@email.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Ling Hong/Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.llnl.gov		
	Plate: ILCLM1838 row: n column: 06		
	High quality sequence steps: 688.		
FEATURES	Location/Qualifiers		
source	1..1092		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:502097"		
	/clone_lib="NIH_MGC_17"		
	/tissue_type="rhadomyosarcoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: muscle; Vector: pOTf7; Site_1: EcoRI;		
	Site_2: XhoI; cDNA made by oligo-dT priming.		
	Directionally cloned into EcoRI/XhoI sites using the		
	following 5' adaptor: GGCAGCAG(C) size selected >500bp		
	for average insert size 1.8kb. Library constructed by		
	Ling Hong in the laboratory of Gerald M. Rubin (University		
	of California, Berkeley) using ZAP-cDNA synthesis kit		
	(Stratagene) and Superscript II RT (Life Technologies)."		
BASE COUNT	346 a 326 c 259 g 161 t		
ORIGIN			

## Alignment Scores:

Pred. NO.: 1.62e-65 Length: 1092  
 Score: 695.50 Matches: 160  
 Percent Similarity: 64.40% Conservative: 39  
 Best Local Similarity: 51.78% Mismatches: 90  
 Query Match: 36.38% Indels: 22  
 DB: 13 Gaps: 4

US-09-908-988b-2 (1-366) x B114696 (1-1092)

```

QY 1 MetasnphethrValGlyPheLysProLeuLeuGlyAspAlaHisasnMetaspasnLeu 20
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 134 ATGAGCGCATCTCGAATTAACAAATCTTTTCCAAAGACAGACAGACATGGATTAACCTTA 193
QY 21 GluysglnLeuIleCysProIleCysLeuGluMetPheSerLysProValIleLeu 40
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 194 GAGAGAGACATCATCTGCTCCATCTGCTTAGAGATGTTACGAAACCTGTGATTC 253
QY 41 ProCysglnHisasnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 254 CCTTGTAGCAACAACCTGTGTAGAAATGTGCCAGTATATTTCAGGCCCTTAACCCG 313
QY 61 LeuTrpInserArgLysSerThrValSerSerGlyGlyArgPheArgCysProSer 80
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 314 TATTTGCCACAAAGAGAGATACCAACCATGGCATTCAGGGGCCGATTCGGCTGATCC 373
QY 81 CysArgHisGlnValValLeuAspArgHisGlnValYrGlyLeuGlnArgAsnLeu 100
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 374 TGTAGACATAGAGGTGTTGCGATAGACATGGGATATGTGACTTCAAGAGAACTGCTG 433
QY 101 ValGlu--AsnIleIleAspIleTyrLysGlnGluSerSerArgProLeuHisAlaLysA 120
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 434 GTGGAGCACTATCATTTGACATTCACAGCGAGGATCCACAGCCCA-----GAAAGA 487
QY 120 LagluGlnHisLeuMetCysGluGlnHisGluAspGluLysIleasnIleTyrCysLeu 140
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 488 AATCCAGCAGCC-ATGTGGAGAGAACATGAGAGAGAGAGAGCATCACTACTGCTGGA 546
QY 140 erCysGlnVal--ProThrCysSerLeuCysLysValPheGlnAlaHisLysAspCysGlu 159
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 547 ACTGCGCAAGTACCAACCTCTCTGTGCAAGGTGA--GGTCACACAAAGACTGCCAG 605
QY 160 ValAla--ProLeuProThrIleTyrLysArgGlnLysSerGluLeuSer--AspGlyIleA 179
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 606 GGGCGTCACCATATCATCTGTACAGAGACAGAGAGTGTAGCTCATGTATGTCATCG 665
QY 179 IametLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGlnGluValC 199
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 666 GCATACTCTGTGGCAGACAGACAGAGTACAGGAGTATCAAGCACTGGAAACACCT 725
QY 199 ysgIlnThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlu 219
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 726 GCAAACTATCGAGGCAATGTGCCGGAAACACCGA-----ACACGAGACCTGA 773
QY 219 hrLeuCysAlaValLeuGlnGluArgLysGlyGluLeuLeuGlnAlaLeuAlaArgLug 239
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 774 CGGAGACAGCGCTGACTACTGGAAGAGATCGAGCCGACAGCAACAAGACCCGCGC 833
QY 239 lngluGlnLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyAspHisLeuGlnG 259
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 834 AGCCACACCCCAACCCACAAACAAAGAGAGAGAGCGC-----GACCACACAAAC 884
QY 259 LysSerLysLeuValGlnSerAlaIleGlnSerMetGlnGluProGlnMetAlaLeu 279
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 885 AGGCAAGTCCGACCTAGAGC-----GACCGTCCGACAAAGAGCGAGAA 926
QY 279 yrLeuGlnGlnAlaLysGlnLeuIleAsnLysValGlyLametSerLysValGlnLeuA 299
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 927 ACAGAAACAGACCCAGACAGCGGCAACAGCAGCAACGTAACGTCACACCGCCACACCTGC 966
QY 299 lagLysArgProGluPro 304
    |||...:|||||
  
```

Db 987 ACACCCACCCGAGCCCA 1003  
 RESULT 19  
 B114212  
 LOCUS  
 DEFINITION  
 B114212  
 mRNA sequence.  
 602862572EP1 NIH\_MGC\_17 Homo sapiens CDNA clone IMAGE:5021825 5',  
 B114212  
 B114212.1 GI:14565113  
 EST.  
 VERSION  
 B114212.1  
 KEYWORDS  
 EST.  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE  
 1 (bases 1 to 776)  
 NIH-MGC http://mgc.ncl.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs+email.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM1840 row: P column: 18  
 High quality sequence stop: 650.  
 Location/Qualifiers  
 1..776  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5021825"  
 /clone\_lib="NIH\_MGC\_17"  
 /tissue\_type="rhabdomyosarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: muscle; Vector: pOTB7; Site\_1: EcoRI;  
 Site\_2: XhoI; CDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(5). Size-selected >500bp  
 for average insert size 1.8kb. Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-CDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT  
 226 a 210 c 190 g 150 t  
 ORIGIN

## Alignment Scores:

Pred. NO.: 1.76e-65 Length: 776  
 Score: 693.00 Matches: 141  
 Percent Similarity: 77.57% Conservative: 25  
 Best Local Similarity: 65.89% Mismatches: 41  
 Query Match: 36.24% Indels: 7  
 DB: 13 Gaps: 1

US-09-908-988b-2 (1-366) x B114212 (1-776)

```

QY 1 MetasnphethrValGlyPheLysProLeuLeuGlyAspAlaHisasnMetaspasnLeu 20
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 134 ATGAGCGCATCTCGAATTAACAAATCTTTTCCAAAGACAGACAGATGGATTAACCTTA 193
QY 21 GluysglnLeuIleCysProIleCysLeuGluMetPheSerLysProValIleLeu 40
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 194 GAGAGAGACATCATCTGCTCCATCTGCTTAGAGATGTTACGAAACCTGTGATTC 253
QY 41 ProCysglnHisasnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 254 CCTTGTAGCAACAACCTGTGTAGAAATGTGCCAGTATATTTCAGGCCCTTAACCCG 313
QY 61 LeuTrpInserArgLysSerThrValSerSerGlyGlyArgPheArgCysProSer 80
    |||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:|||||
Db 314 TATTTGCCACAAAGAGAGATACCAACCATGGCATTCAGGGGCCGATTCGGCTGCAATCC 373
  
```



•  
•  
•  
•

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 09:04:09 ; Search time 2053 Seconds  
(without alignments)  
11288.720 Million cell updates/sec

Title: US-09-908-988B-1  
Perfect score: 1331  
Sequence: 1 aagagatgtagacagagtglt.....ataaagacacgaagtcc 1431

Scoring table: IDENTITY-NUC  
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

TEST: 7  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_huv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	538.2	37.6	638	9	AA800245
2	508.2	35.5	608	12	BF284900
3	482.4	33.7	485	11	AK019655
4	443	31.0	555	12	BF855618
5	439.2	30.7	516	13	BM030114
6	420.8	29.4	461	9	AV006036

*166 over 100  
the prob of  
all the hits  
and hits 2-5*

C	7	416.2	29.1	512	12	BF855619	BF855619 RCO-FN020
	8	416.2	29.1	595	13	BJ071661	BJ071661
	9	413.6	28.9	432	10	BB732046	BB732046
C	10	403.2	28.2	504	9	A1712700	A1712700 UI-R-AFL-
	11	361.2	25.2	392	9	AV006120	AV006120
C	12	360	25.2	456	10	BE099923	BE099923 UI-R-BL1-
	13	350.2	24.5	912	13	B1758535	B1758535 603022861
	14	345.4	24.1	658	14	BM726338	BM726338 UI-E-EJO-
	15	342.6	23.9	793	13	B1754358	B1754358 60302606
	16	341.8	23.9	378	10	AW918753	AW918753 EST350057
C	17	341.4	23.9	440	10	BC375657	BC375657 UI-R-CSO-
	18	340.6	23.8	441	12	BG379943	BG379943 UI-R-CSO-
C	19	333	23.3	868	12	BM48618	BM48618 pgm2n-PK0
	20	327.8	22.9	495	13	BC761218	BC761218 602718564
	21	316	22.1	1105	12	BM483766	BM483766 56372 MAR
	22	312.8	21.9	373	10	BM520266	BM520266
C	23	303.2	21.2	694	10	BE095943	BE095943 UI-R-BUO-
	24	289.2	20.2	386	10	BE095943	BE095943 UI-R-BUO-
	25	289	20.2	316	10	BB140247	BB140247
	26	284	19.8	721	12	BG674823	BG674823 602620959
	27	271.2	19.0	790	12	BF679996	BF679996 602745109
	28	260	18.2	682	12	BF309583	BF309583 601891901
	29	259.6	18.1	613	13	BJ033623	BJ033623
	30	258.6	18.1	1161	12	EG163322	EG163322
	31	258.4	18.1	450	13	BM087922	BM087922 501235 MA
	32	255.6	17.9	983	12	BE205857	BE205857 60168260
	33	249.4	17.4	291	12	BF554230	BF554230 UI-R-CO-h
	34	246.8	17.2	1092	13	BT114696	BT114696 602861723
	35	244.2	17.1	475	9	AA428229	AA428229 zvr3g06.r
	36	243.8	17.0	1081	12	BF309583	BF309583 601891901
	37	240.2	16.8	1065	17	CNS04GLR	AL286800 Tetradon
C	38	239.4	16.7	420	9	AA447961	AA447961 zvr3g08.s
	39	239.2	16.7	581	12	BF829678	BF829678 MR2-HN003
C	40	237.8	16.6	329	9	A1763576	A1763576 UI-R-YO-a
	41	237.6	16.6	776	13	BT114212	BT114212 602862572
	42	235.6	16.5	672	10	BB662321	BB662321
	43	234.6	16.4	464	9	A1644642	A1644642 vwr79p06.x
	44	233.8	16.3	573	9	AA443443	AA443443 zvr4c05.t
	45	233.8	16.3	628	13	BJ059953	BJ059953

## ALIGNMENTS

RESULT 1  
LOCUS AA800245/c 638 bp mRNA linear EST 30-APR-1998  
DEFINITION EST189742 Normalized rat heart, Bento Soares Rattus sp. cDNA clone  
RHEM89 3' end, mRNA sequence.

ACCESSION AA800245  
VERSION AA800245.1 GI:2863200

KEYWORDS EST.  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryota; Metazoa; Chordata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 638)

AUTHORS Lee N.H., Glodok A., Chandra I., Mason T.M., Quackenbush J.,  
Kerlavagne A.R. and Adams M.D.  
Rat Genome Project: Generation of a Rat EST (RES) Catalog & Rat  
Gene Index

JOURNAL Unpublished (1998)

COMMENT The Institute for Genomic Research  
9712, Medical Center Drive, Rockville, MD 20850, USA  
Tel: (301)-838-3529  
Fax: (301)-838-0208  
Email: nhlee@tigr.org  
Seq primer: M13-21.  
Location/Qualifiers

FEATURES  
Source 1..638  
/organism="Rattus sp."

```

/db_xref="ATCC (inhost):2007238"
/db_xref="taxon:10118"
/clone="RHEM89"
/clone_lib="Normalized rat heart, Bento Soares"
/Note="Organ: heart; Vector: pT37Pac; Site_1: EcoRI;
Site_2: NotI"
BASE COUNT      105 a      221 c      165 g      147 t
ORIGIN
Query Match      37.6%; Score 538.2; DB 9; Length 638;
Best Local Similarity 91.7%; Pred. No. 2.4e-107;
Matches 595; Conservative 0; Mismatches 43; Indels 11; Gaps 2;

QY 783 GGAGAGAGTGTCCAGACATTTGAGAGACAAGCCGACAGACAGACACTGTAAACCA 842
DB 638 GGAGAGAGTGTCCAGACATTTGAGAGACAAGCCGACAGACAGACACTGTAAACCA 579
QY 843 GAGGTTCCAGACCCCTGTGGCGGTTTGGAGAGAGCGCAAGGGGCAACTGCTTCAAGCACT 902
DB 578 GAGGTTCCAGACCCCTGTGGCGGTTTGGAGAGAGCGCAAGGGGCAACTGCTTCAAGCGCT 519
QY 903 GCGCGGAGAGAGAGAGAGAGTTCAGCGCGTGGGGGCTTCATCCGACATCGAGAGA 962
DB 518 GCGTCGAGTGAAGAGAGAGAGTTCAGCGCGTGGGGGCTTCATCCGACATCGAGAGA 459
QY 963 CCACCTTGAGAGGCTCTCTAAAGCTGTGAGTCCGACATCCATCGAGAGCGCGCA 1022
DB 458 CCACCTTGAGAGGCTCTCTCTAAAGCTGTGAGTCCGACATCCATCGAGAGCGCGCA 399
QY 1023 GATGGCTCTACCTCCAGCAGAGCAAGAGAGTGTATCAACAGTGGGGCAATGTGCA 1082
DB 398 GATGGCTCTACCTCCAGCAGAGCAAGAGAGTGTATCAACAGTGGGGCAATGTGCA 339
QY 1083 GGTGAGACTGTGAGAGAGCGCGGAGCCAGGCTATGAGAGCATGAGCAATCTGTGTAG 1142
DB 338 GGTGAGACTGTGAGAGAGCGCGGAGCCAGGCTATGAGAGCATGAGCAATCTGTGTAG 279
QY 1143 CTTGAGACAGTGTGCGCAATGTTGCGAATCGACTTCCAGCGCGCGCTGGGGA 1202
DB 278 CTTGAGACAGTGTGCGCAATGTTGCGAATCGACTTCCAGCGCGCGCTGGGGA 219
QY 1203 TGAAGAGATGACACATGCTTGGATGGGAGAGAGGCAATGGCGGCTGGAGAGAGA 1262
DB 218 TGAAGAGATGACACATGCTTGGATGGGAGAGAGGCAATGGCGGCTGGAGAGAGA 159
QY 1263 GCGGCTGTGAGAGTGTGCGAAGGCTCAGCTGACCTGACCTGTGATCCAGAGCGCAG 1322
DB 158 GCGGCTGTGAGAGTGTGCGAAGGCTCAGCTGACCTGACCTGTGATCCAGAGCGCAG 105
QY 1323 ACCGAGAGCGGAGGCAAGGATGCTGAGATCTGCGGAGAGAGACCGCGCCACCAAGC 1382
DB 104 GCC-----CGGAGGCCAAGGAGATCTGAGATGTCGCGAGAGACCATCTACCCATCC 50
QY 1383 TCGGCTTCCGCGCCCGGAGAGTTCATTAAGAGACTCAAGTGTCCC 1431
DB 49 TCGGCTTCCGCGCCCGGAGAGTTCATTAAGAGACTCAAGTGTCCC 1

RESULT 2
BF284900/c      608 bp      mRNA      linear      EST 28-NOV-2000
LOCUS          EST149491 Rat Gene Index, normalized rat, Rattus norvegicus cdna
DEFINITION     Rattus norvegicus cdna clone RGIEX57 3' sequence, mRNA sequence.
ACCESSION      BF284900
VERSION        BF284900.1 GI:11215970
KEYWORDS       EST.
SOURCE         Norway rat.
ORGANISM       Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE      1 (bases 1 to 608)
AUTHORS        Malek,R.L., Cho,J., Lee,Y., Karamycheva,S., Parvizli,B., Pertea,G.,

```

```

TITLE          Sultana,R., Tsai,J., White,J., Quackenbush,J. and Lee,N.H.
JOURNAL        Generation of ESTs from Normalized Rat Embryo, Bento Soares
COMMENT        Unpublished (2000)
                Other ESTs: EST350057
                Contact: Lee, NH
                The Institute for Genomic Research
                9712, Medical Center Drive, Rockville, MD 20850, USA
                Tel: (301)-838-3529
                Fax: (301)-838-0208
                Email: nhlee@tigr.org
                This clone is available through the ATCC, contact the ATCC
                tel#703-365-2700 for further information.
FEATURES
Source
/db_xref="taxon:10116"
/clone="RGIEX57"
/clone_lib="Rat Gene Index, normalized rat, Rattus
norvegicus cdna"
/tissue_type="mixed tissue"
/lab_host="DH5-alpha"
/Note="Vector: pT37Pac; Site_1: EcoRI; Site_2: NotI;
Combination of ROV, RBR, RRI, RLI, RPL, RLU, REM, RMU, RSP
RHE, RPC, RPN"
BASE COUNT      101 a      207 c      162 g      138 t
ORIGIN
Query Match      35.5%; Score 508.2; DB 12; Length 608;
Best Local Similarity 91.3%; Pred. No. 8.5e-101;
Matches 565; Conservative 0; Mismatches 43; Indels 11; Gaps 2;

QY 813 CAGCGGAGACAGAGACACTGTTAAACAGAGTTCGAGACCTGTGCGGTTTGA 872
DB 608 CAGCGGAGACAGAGACACTGTTAAACAGAGTTCGAGACCTGTGCGGTTTGA 549
QY 873 GGAGCGCAAGGGGCAATGCTTCAAGCAGTGGCCGGGAGAGAGAGAGTGCACAGC 932
DB 548 GGAGCGCAAGGGGCAATGCTTCAAGCAGTGGCCGGGAGAGAGAGAGTGCACAGC 489
QY 933 CTTGCGGGGCTTATCCGCCAGTACGAGAGACCACTTGGAGGCTCTCAAGCTGTGGA 992
DB 488 CTTGCGGGGCTTATCCGCCAGTACGAGAGACCACTTGGAGGCTCTCTCAAGCTGTGGA 429
QY 993 GTCCGCATCATCATGATGAGAGCGCGAGATGGCTCTACTCTCCAGCGCAAGAGA 1052
DB 428 GTCCGCATCATCATGATGAGAGCGCGAGATGGCTCTACTCTCCAGCGCAAGAGA 369
QY 1053 GCTGATCAACAGAGTGGGCAATGTGCAAGTGTGAGACTGGAAGAGCGCGAGAGG 1112
DB 368 GCTGATCAACAGAGTGGGCAATGTGCAAGTGTGAGACTGGAAGAGCGCGAGAGG 309
QY 1113 CTATGAGAGATGAGACCAATTTCTGTGAGCTGTGAGACAGTGGCGCAATGTTGCAAC 1172
DB 308 CTATGAGAGATGAGACCAATTTCTGTGATGTGAGACAGTGGCGCAATGTTGCGGAC 249
QY 1173 CATGCACTTCCAGCCCGCGGCTGGGATGAAGAGATGATGATGATGATGATGATG 1232
DB 248 CATGCACTTCCAGCCCGCGGCTGGGATGAAGAGATGATGATGATGATGATGATG 189
QY 1233 GGAGAGGCGCAATGCGGCTGGAGAGAGCGGCTGTGAGAGTCCAGAGAGCTTCAAG 1292
DB 188 GGAGAGGCGCAATGCGGCTGGAGAGAGCGGCTGTGAGAGTCCAGAGAGCTTCAAG 135
QY 1293 GCACGTACCCAGCTGTGATCCAGAGGCGACACCGAAGGCGGAGCCCAAGGATGCTGAG 1352
DB 134 GCACGTATCCAGCTGTGATCCAGAGGCGACGCC-----CGGAGGCCAAGGATGCTGAG 80
QY 1353 ATCTGCGCAGAGACACCGCGCCACCAAGCTGGCTTCCCGCGCGGAGAGGTTCTCA 1412
DB 79 ATCTGCGCAGAGACCATGCGGCTTACCATCTGGGCTCCACCCCGAGGATGTTCTCA 20
QY 1413 TAAAGGACTCAAGTGTCCC 1431

```





Db 122 GTGCGGGGCAATGTCAGAGGTGAGCTGCGAGACGGCCGACCGACGCTATGAGAGCATG 181  
 QY 1126 GAGCAATTCCTGTGTGCGGTGAGACGACGTGGCGGAATGTTGGAGACATCGACTTCCAG 1185  
 Db 182 GAGCAATTCCTGTGTGCGGTGAGACGACGTGGCGGAATGTTGGAGACATCGACTTCCAG 241  
 QY 1186 CCGGCGCGCGCTGGGGATGAGAGATGACGACATGCTTTGGATGGGAGGAGGCAAT 1245  
 Db 242 CCGGCGCGCGCTGGGGATGAGAGATGACGACATGCTTTGGATGGGAGGAGGCAAT 301  
 QY 1246 GCGGGGCTGGAGAGAGAGCGCTGAGACGTCCAGAAAGCTCAGGCTGCATGACCCGAC 1305  
 Db 302 GCGGGGCTGGAGAGAGAGCGCTGAGACGTCCAGAAAGCTCAGGCTGCATGACCCGAC 361  
 QY 1306 TCTGATCCAGAGCGCACACCCGAGCGGAGGCCAAGGATGCTGAGAGATGCTGGCAGAGA 1365  
 Db 362 TCTGATCCAGAGCGCACACCCGAGCGGAGGCCAAGGATGCTGAGAGATGCTGGCAGAGA 421  
 QY 1366 CCACGCGCGCACCAAGCTCGGCTCCGCGCCCGGGAGAGTTCTCAATTAAGACTCAAG 1425  
 Db 422 CCACGCGCGCACCAAGCTCGGCTCCGCGCCCGGGAGAGTTCTCAATTAAGACTCAAG 481  
 QY 1426 TGTG 1429  
 Db 482 TGCC 485

RESULT 4  
 BF855618 555 bp mRNA linear EST 16-JAN-2001  
 LOCUS RCO-FN0204-161100-031-d06 FN0204 Homo sapiens CDNA, mRNA sequence.  
 DEFINITION BF855618  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 human.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS  
 1 (bases 1 to 555)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,  
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 CONTACT: Simpson A.J.G.  
 20202663  
 COMMENT  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC0&t2=RC0-FN0204-  
 161100-031-d06&t3=2000-11-16&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 51  
 High quality sequence stop: 120.  
 Location/Qualifiers  
 1..555  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="FN0204"  
 /dev\_stage="Adult"  
 /note="Organ: prostate.normal; Vector: puc18; Site\_1: SmaI  
 ; Site\_2: SmaI; A mini-library was made by cloning  
 products derived from ORESTES PCR (U.S. Letters Patent

## FEATURES

Location/Qualifiers

1..555

organism="Homo sapiens"

db\_xref="taxon:9606"

clone\_lib="FN0204"

dev\_stage="Adult"

note="Organ: prostate.normal; Vector: puc18; Site\_1: SmaI  
 ; Site\_2: SmaI; A mini-library was made by cloning  
 products derived from ORESTES PCR (U.S. Letters Patent

application No. 196/716 - Ludwig Institute for Cancer  
 Research) profiles into the pUC 18 vector. Reverse  
 transcription of tissue mRNA and cDNA amplification were  
 performed under low stringency conditions."  
 BASE COUNT 113 a 149 c 167 g 126 t  
 ORIGIN

Query Match 31.0%; Score 443; DB 12; Length 555;  
 Best Local Similarity 91.3%; Pred. No. 1.5e-86;  
 Matches 470; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 276 CCCCATCTGCTGGAGATGTTCTTCAAGCCGCTGGTATCTTCCCTGCCAACAACCT 335  
 Db 555 CCCCATCTGCTGGAGATGTTCTTCAAGCCGCTGGTATCTTCCCTGCCAACAACCT 496  
 QY 336 GTGCGCAAGTGTGTCACACGACGTTCTCAGGCTCTATCTCTGTGGCATCCCGGG 395  
 Db 495 GTGCGCAAGTGTGTCACACGACGTTCTCAGGCTCTATCTCTGTGGCATCCCGGG 436  
 QY 396 CTCACAAAGGTGCTTTCAGAGAGAGCTTTCGATGCCATCTTGAAGCAGAGTTGT 455  
 Db 435 CTCACCAAGTGTGCTTTCAGAGAGAGCTTTCGATGCCATCTTGAAGCAGAGTTGT 376  
 QY 456 CTTGAGACGAGCATGTGTCATGTCGTCAGGAGCACTGCTAGTGTGAGAGACATCTTGA 515  
 Db 375 CTTGAGACGAGCATGTGTCATGTCGTCAGGAGCACTGCTAGTGTGAGAGACATCTTGA 316  
 QY 516 CATCTACAGCAGAGTCTCCCGCCACTCCAGCCCAAGGCTGAAACGACCTCATGTG 575  
 Db 315 CATCTACAGCAGAGTCTCCCGCCACTCCAGCCCAAGGCTGAAACGACCTCATGTG 256  
 QY 576 TGAGAGCAGAGAGAGAGAGATCAATCTACTGCTGAGCTGCGAGGTGCCACCTG 635  
 Db 255 CGAGAGCAGAGAGAGAGAGATCAATCTACTGCTGAGCTGCGAGGTGCCACCTG 196  
 QY 636 CTCTCTCTGCAAGTTTGGCGCCGCCAAGAGACTGAGGTGGCCCTGTCACCATCAT 695  
 Db 195 CTCTCTCTGCAAGTTTGGCGCCGCCAAGAGACTGAGGTGGCCCTGTCACCATCAT 136  
 QY 696 TTTCACAAAGCCCAAGAGTGTGAGCTGAGCAGATGGCATGCGATGCTGTGGCGGCAATGA 755  
 Db 135 TTTCACAAAGCCCAAGAGTGTGAGCTGAGCAGATGGCATGCGATGCTGTGGCGGCAATGA 76  
 QY 756 CCGTGTGACAGCAGTGTACCCACCATGATGAGAGAG 790  
 Db 75 CCGTGTGACAGCAGTGTACCCACCATGATGAGAGAG 41

RESULT 5  
 BM030114 516 bp mRNA linear EST 05-NOV-2001  
 LOCUS 488634 MARC 2B0V Bos taurus CDNA 5', mRNA sequence.  
 DEFINITION BM030114  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.  
 1 (bases 1 to 516)  
 Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrnkung,S.C., Bennett,  
 G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,  
 Perteau,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and  
 Keele,J.W.  
 Sequence evaluation of four pooled-tissue normalized bovine cDNA  
 libraries and construction of a gene index for cattle  
 Genome Res. 11 (4), 626-630 (2001)  
 21180013  
 CONTACT: Smith TPL  
 USDA, ARS US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA

TITLE  
 JOURNAL  
 MEDLINE  
 COMMENT



Accession	Sequence	Position
Db	GGATGACGACATGCGTTTGATGGGAGGAGGCAATGCGGGGCTGGAGAGAGCGGCT	298
QY	GGAGCTGGCAGAGGCTCAGGCTGCACGTGACCCGACTTGATCCAGAGCGCACACCGA	1328
Db	GGAGCTGCCAGAGGCTCAGGCTGCACGTGACCCGACTTGATCCAGAGCGCACACCGA	358
QY	AGCGGAGCCAAAGGATCTGAGATCTGCAGAGAGACCAGCGCGCCAAAGCTCGGCT	1388
Db	AGCGGAGCCAAAGGATCTGAGATCTGCAGAGAGACCAGCGCGCCAAAGCTCGGCT	418
QY	TCCCGCCCCGGGAAGTTCTCAATAAAGAACTCAAGTGTCC	1431
Db	TCCCGCCCCGGGAAGTTCTCAATAAAGAACTCAAGTGTCC	461

RESULT 7	LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM
BF855619/c	BF855619	512 bp	MRNA	linear	EST 16-JAN-2001	
	RC0-FN0204-161100-031-e06	FN0204	Homo sapiens	CDNA, mRNA sequence.		
	BF855619.1	GI:12243363	EST.			
	human.					
	Homo sapiens					

REFERENCE  
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 512)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.

TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G.

Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP  
Brazil

Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
 This sequence was derived from the FAPESP/ICR Human Cancer Genome project. This entry can be seen in the following URL  
<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC06t2=RC0-FN0204>  
 161100-031-e066t3-2000-11-166t4-1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 42  
 High quality sequence stop: 512.

```

SOURCE
LOCATION/Qualifiers
1. 512
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="FND204"
/dev_stage="Adult"
/note="Organ: prostate.normal; Vector: puc18; Site:1: Sma1
: Site:2: Sma1; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
104 a 138 c 157 g 113 t

```

Query Match	29.18;	Score 416.2;	DB 12;	Length 512;
Best Local Similarity	91.08;	Pred. No. 1e-80;		
Matches 465;	Conservative 0;	Mismatches 43;	Indels 3;	Gaps 2

QY	276	CCCCATCTGCGTGGAGATGTTCTCCAGGCCCGTGGTATCTTGGCCCTGCCAACAAACCT	335
Db	512	CCCCATCTGCGTGGAGATGTTCTCCAAACAGAGTGATCTCTGGCCCGCCAAACAACT	453
QY	336	GTCGCCAAGCTGTGCCAACACAGCTCTTCCAGGCCCTTAATCTCTGTGGCAATCCGGGG	395
Db	452	GTCGCCCAATGTGCCAACACAGCTTCTCCAGGCCCTGAATCTCTATGGCAGTCCGGGG	393
QY	396	CTCCACAAGGTTCTCTTCAGAGGAGTGTCCGATGCCATCTTTGTAGGCACGAGTTGT	455
Db	392	CTCCACACACTGTGTCTTTCAGAGGAGCCGTTTCCCTGTGCCCATGTCTGACGAGGATGAGTTGT	333
QY	456	CCTGGACAGGCATGTGTCTATGGCCTGCACGCCGAACCTGTAGTGAGAAACATCATTTGA	515
Db	332	CCTGGACACACGCGTGTCTACGGCCTGCACGAAACCTGTAGTGGAGAACTATACGA	273
QY	516	CATCTCAAGCAGGAGTTCCTCCCGGCGCATCGACGCCGAAGGCTGAACAGCACTCATGTG	575
Db	272	CATTTTACAAGCAGGAGTCACTCCAGGCGGTGCATCTCAAGGCTGACACAGCACTCATGTG	213
QY	576	TGAGGAGCAGGAGGAGAGATCAACATCTACTCCCTGAGCTGGAGGTGCCCACTG	635
Db	212	CGAGGAGCATTGAAGAAGAGATCAATTTTACTGCTCGAGCTGGAGGTGCCCACTG	153
QY	636	CTCTCTGTCGCAAGTGTTCGGGCGCCACAAAGACACTGTGAGGTGGCCCCCTCTGTGCCACCAT	695
Db	152	CTCTCTCTGTCGCAAGTCTTCGGGTGCCACAAAGACACTGTGAGGTGGCCCCCACTGTGCCACCAT	93
QY	696	TTACCAACG--CCAGAAAGTGTAGCTGACGATGGCATGGCATGCTGTGGGGGGCAAT	753
Db	92	TTACTTAAGCGCCCCGAAAGATGAGCTCAGCGATGGATGGCATGTGTGGCAGGCAAT	33
QY	754	GACCGTGTGCAAGCA--GTGATACACCAAGATG 783	
Db	32	GACCCGTCGCAAGCATGTGATCACACCGATG 2	

RESULT	8
Bj071661	
LOCUS	Bj071661 595 bp mRNA linear EST 11-DEC-2001
DEFINITION	Bj071661 NIH Mochi normalized Xenopus tailbud library xenopus
ACCESSION	laevis cDNA clone Xl095b10 5', mRNA sequence.
VERSION	Bj071661
KEYWORDS	Bj071661.1 GI:17501850
SOURCE	EST:
ORGANISM	African clawed frog. Xenopus laevis

REFERENCE	1 (bases 1 to 595)
AUTHORS	Kitayama, A., Terasaka, C., Mochi, M., Ueno, N., Shin-i, T. and Kohara, Y.
TITLE	Expressed genes in X. laevis embryo
JOURNAL	Unpublished (2001)
COMMENT	Contact: Tadasu Shin-i

FEATURES  
source  
1.595  
Location/Qualifiers  
Email: [tshini@genes.nig.ac.jp](mailto:tshini@genes.nig.ac.jp)  
Fax: 81-559-81-6855  
Tel: 81-559-81-6856  
Iida, Iada, Mishima, Shizuoka 411-8540, Japan

```

/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL095b10"
/clone_lib="NRB Mochil normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/ncbi_vector="pBSR3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dt primed and directionally cloned; Staging

```



Db 61 TCACAGAGTCGGGGCATGTGCAAGTGGAGCTGGCAGACCGCCGACCAAGCTATG 120  
 QY 1118 AGAGATGAGCAATTTCTGTGAGCGTGGAGCAGCTGGCCGAATGTTGCAACCATCG 1177  
 Db 121 AGACATGAGCAATTTCTGTGAGCGTGGAGCAGCTGGCCGAATGTTGCAACCATCG 180  
 QY 1178 ACTTCACGCGCGCGCCGCTGGATGAGAGATGACGATGCTTGGATGGGAGG 1237  
 Db 181 ACTTCACGCGCGCGCCGCTGGATGAGAGATGACGATGCTTGGATGGGAGG 240  
 QY 1238 AGGCAATGCGGGGCTGAGAGAGAGCGGCTGACGTGCAGAGGCTCAGGCTGACT 1297  
 Db 241 AGGCAATGCGGGGCTGAGAGAGAGCGGCTGACGTGCAGAGGCTCAGGCTGACT 300  
 QY 1298 GACCCGATCTGATGCAAGCGCACACCCGGAAGCGGAGCCAGAGGATGCTGAGATCTG 1357  
 Db 301 GACCCGATCTGATGCAAGCGCACACCCGGAAGCGGAGCCAGAGGATGCTGAGATCTG 360  
 QY 1358 CGCAGAGACACCGCGCCACCAAGCTCGGCTCCGCGCGCGGGAAGTTCTCAATAAG 1417  
 Db 361 CGCAGAGACACCGCGCCACCAAGCTCGGCTCCGCGCGCGGGAAGTTCTCAATAAG 420  
 QY 1418 GACTCAAGTGC 1429  
 Db 421 GACTCAAGTGC 432

RESULT 10  
 AT172700/c 504 bp mRNA linear EST 08-JUN-1999  
 LOCUS AT172700  
 DEFINITION UI-R-AFL1-aa-e-02-0-UI.s1 UI-R-AFL1 Rattus norvegicus cDNA clone  
 VERSION AT172700  
 KEYWORDS AT172700.1 GI:5016500  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 504)  
 Bonaudo, M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 97044477

JOURNAL CONTACT: Soares, MB  
 MEDLINE Program for Rat Gene Discovery and Mapping  
 COMMENT University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a Bonafide poly A  
 and the oligo-dT track served to verify it as a clone from the  
 normalized AV canal at 15 dpc library cDNA library Preparation:  
 M.B. Soares Lab Clone distribution: clones will be available  
 through Research Genetics (www.resgen.com)  
 Seq primer: M13 Forward  
 POLYA=Yes.

# FEATURES

1..504  
 Location/Qualifiers  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-AFL1-aa-e-02-0-UI"  
 /clone\_11b="UI-R-AFL1"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-AFL1  
 library is a normalized library constructed from 15 dpc  
 rat atrioventricular (AV) canal. The tag is a string of 5  
 nucleotides present between the Not I site and the  
 oligo-dT track. The library was constructed as described  
 by Bonaudo, Lennon and Soares, Genome Research 6: 791-806  
 , 1996. Tissue provided by Jim Lin, Department of Biology,  
 University of Iowa.  
 TAG-LIB=UI-R-AFL1  
 TAG-TISSUE=AV canal at 15 dpc  
 TAG-SEQ=GAAGC

Query Match 28.2%; Score 403.2; DB: 9; Length 504;  
 Best Local Similarity 91.2%; Pred. No. 7.1e-76;  
 Matches 454; Conservative 0; Mismatches 33; Indels 11; Gaps 2;

QY 934 GTGCGGGGCTCATCCGCCAGTACGAGACCACTTGGAGGCTCTCAAGAGCTGGAG 993  
 Db 504 GTGCGTAGCTCATCCGCCAGTACGAGACCACTTGGAGGCTCTCTCAAGAGCTGGAG 445  
 QY 994 TCCGCCATCCAGTATGAGGAGAGCCGAGATGGCTCTCTACCTCCAGCAGCAAGAG 1053  
 Db 444 TCCGCCATCCAGTATGAGGAGAGCCGAGATGGCTCTCTACCTCCAGCAGCAAGAG 385  
 QY 1054 CTGATCAACAAGTTCGGGGGCAATGTCGAAGTGGAGCTGGCAGAGACCGCCGAGCAGG 1113  
 Db 384 CTGATCAACAAGTTCGGGGGCAATGTCGAAGTGGAGCTGGCAGAGACCGCCGAGCAGG 325  
 QY 1114 TATGAGAGCATGAGCAATTTCTGTGAGCGTGGAGCAGCAGTGGCCGAATGTTGCAAGC 1173  
 Db 324 TATGAGAGCATGAGCAATTTCTGTGAGCGTGGAGCAGCAGTGGCCGAATGTTGCAAGC 265  
 QY 1174 ATGCACTTCCAGCGCGGGCGCGCTGGGATGAAGAGATGACGATGCGCTTGGATGGG 1233  
 Db 264 ATGCACTTCCAGCGCGGGCGCGCTGGGATGAAGAGATGACGATGCGCTTGGATGGG 205  
 QY 1234 GAGGAGGCAATCGGGGCTGAGAGAGAGCGGCTGAGCAGTGCAGAGGCTCAGGCTG 1293  
 Db 204 GAGGAGGCAATCGGGGCTGAGAGAGAGCGGCTGAGCAGTGCAGAGGCTCAGGCTG 151  
 QY 1294 CACTGACCCGACTCTGATTCAGAGCGCACGCCGAGCGGAGGAGGATGCTTGCAT 1353  
 Db 150 CACTGATCCGACTCTGATTCAGAGCGCACGCCGAGCGGAGGATGCTTGCAT 96  
 QY 1354 TCTGCGCAGAGACACCGCGCCACCAAGCTGGCTTCCCGCCCGGAGAGGTTCTCAAT 1413  
 Db 95 TCTGCGCAGAGACACCATCGCGCTACCATCTCGGCTCCACCCAGGAGATGTTCTCAAT 36  
 QY 1414 AAAGGACCTCAAGTGTCC 1431  
 Db 35 AAAGGACCTCAAGTGTCC 18

RESULT 11  
 AV006120 392 bp mRNA linear EST 25-AUG-1999  
 LOCUS AV006120  
 DEFINITION AV006120 Mus musculus C57BL/6J heart Mus musculus cDNA clone  
 VERSION AV006120  
 KEYWORDS AV006120.1 GI:4782970  
 SOURCE EST.  
 ORGANISM house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 392)

REFERENCE Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,  
 Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara,  
 'A., Hayatsu, N., Horii, F., Ishikawa, T., Itoh, M., Iwama, M., Kawai, J.,  
 Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuma, H., Oda, H., Owa, C.,  
 Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara



QY	1100	GGCCGAGACGACGATGAGACATGAGACCAATTCCTCTGTGACCGCTGGAGCAGCAGTGGCCG	1159
Db	336	GGCCGAGACGACGATGAGACATGAGACCAATTCCTCTGTGATCGGGAGCAGCTGGCCG	277
QY	1160	AAATGTGGCAACATCGACTTCGACGCCGGCGCGCTGGGGATGAAGGATGAGACA	1219
Db	276	AAATGTGGGGACCATCGACTTCGACGCCGGCGCGCTCTGGGGATGAAGGATGAGACGAG	217
QY	1220	TGGCTTTGGATGGGAGGAGGCGCATGCGCGGCGCTGGAGGAGGAGGCGCTGGAGCTGGCCAG	1279
Db	216	TGACTTTGGAGGGGAAAGAGGGCAACACAGGGCTGGAGGAGGAGGCGCTGGACGGCCAG	157
QY	1280	AAGGCTCAGGCGCTGCACATGACCCGACTCTGATCCAGAGCGCACACCCCGAAGCGGGAGCCA	1339
Db	156	AAGGTT-----TGACATGATCCGACTCTGATCCAGAGCGCACGCCG----CGGGAGCCA	108
QY	1340	AGGAGTGTGAGGATCTGCGCAGAGAACCCAGCGCGCACCAACCTGAGGCTTCCCGCCCGG	1399
Db	107	AGGAGTGTGAGGATCTGCGCAGAGAACCATCGCGCTAACCCATCTGGCGCTCCACCCCGAG	48
QY	1400	GGAAGTCTCTCAATAAAGACTCAAGTCTC	1429
Db	47	GGATGGTTCTCATTAAGACTCATGTGTC	18

RESULT_13				
B1758535				
LOCUS	912 bp	mRNA	linear	EST-25-SEP-2001
DEFINITION	603020286.F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5193326 5' , mRNA sequence.			
ACCESSION	B1758535			
VERSION	B1758535.1 GI:15750113			
KEYWORDS	EST .			
SOURCE	human.			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 913)				
NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .				
National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)				
Contact: Robert Strausberg Ph.D.				

FEATURES

source	
	Location/Qualifiers
	1. .912
	High quality sequence start: 25
	High quality sequence stop: 890.
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>
	CDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
	DNA Sequencing by: Incyte Genomics, Inc.
	Tissue Procurement: Life Technologies, Inc.
	CDNA Library Preparation: Life Technologies, Inc.

FEATURES	location/Qualifiers
SOURCE	1. .912
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone IMAGE:5193326"
	/clone_11b="NIH_MGC_114"
	/lab_host="DH10B"
	/note="Organ: brain; Vector: PCMV-SPORT6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-77 yo. Library is oligo-AT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (University). Research Genetics tracking code 019. Note: this is a NIH_MGC library."
BASE COUNT	236 a 232 c 266 g 178 t
ORIGIN	

Query Match	24.5%	Score 350.2;	DB 13;	Length 912;
Best Local Similarity	70.1%;	Pred. No. 3.6e-66;		
Matches 515; Conservative	0;	Mismatches 213;	Indels 7;	Gaps 3

QY	226	CTGCTTAGGGGATATCGGCACCAACATGAGACAACTTGGAGAAAGACAGTCAATTGGCCCATCTGC	285
Db	167	CTGATCCAGGATGGGAATCCCATGGGAGAACTTGGAAACACAGGTGATCTGCTCATCTTC	226
QY	286	CTGGAGATGTTCTCCAGGCCCGTGGATGCTTTCCTGGCCAAACACAACTGTGGCCGACAG	345
Db	227	CTGGAGATGTTTAAACCAAGCCATGTGATCTTTCCTGGCCAAACACAACTGTGGCCGACAG	286
QY	346	TGTGGCAACGACGCTCTTCCAGGCTCTTAATCCTCTGTGGCAATCCCGGGGCTCCACACG	405
Db	287	TGTGCCAAAGACATCTTCCAGGCTGTCAAAATCCCTACTGTGACACCGCGGGACGCTCAGTG	346
QY	406	GTGTCTCTGGAGGAGGTTTCCGATCCCATGTTGTGTAAGGCACAGAGGTTTCTGTGACAGG	465
Db	347	TCCATGTCTGGAGGCGCGTTTCCGCTGCCCCACCTGCGCCACAGAGGTGATCGATGCTCT	406
QY	466	CATGTGTCTATGCGCTGCAGCGGAACTGCTAGTGGAAACATCATATTACATCTACAG	525
Db	407	CACGGAGTGTACGCGCTGCAGAGAAACCTGCTGTGGTGGAAACATCATCTACATCTACAA	466
QY	526	CAGGAGTCTCCCCGGCCACTGCACGCGCAAGGCTGAACAGCACCTCATGTGTGAGGAGCAC	585
Db	467	CAGGAGTGTCTCAGTGGCGGCGCTGCAGAG--GGCAGTCACCCCATGTCGAAGGAGCAC	523
QY	586	GAGGACGAGAAATCAACATCTACTGCGCTGAGCTGCGAGTGGGCCCCACCTGCTCTCTGC	645
Db	524	GAAAGTGAAGAAATTAACATCTACTGTCTCAGCTGTGAAGTGGCCACCTGCTCTCAATGTCG	583
QY	646	AAGTTTTTCGGCGCCACAGAGCATGTGAGGTGGCCCCCTGTGCCACCATTTACAAACGCG	705
Db	584	AAGGTGTTTGGGATCCACAAAGGCTTGCAGGGTGGGCCCATTTGCAAGTGTCTTCCAGGGA	643
QY	706	CAGAGAGTGAAGCTAAGCGATGCGATGCTGTGTGGCGGGCAATGACCGTGTGCAG	765
Db	644	CAAAAGACTGAACCTGAATACTGTATCTTCATGTGCTGTGGGGGGAATGACCGTGTGCAG	703
QY	766	GCAGATGATCACCCAGATGAGGAGTGTGCCAGACCATTTGAGGA--CAACACCGGACAGCA	824
Db	704	AOCATCATCTACTCAGACTGGAGGATTTCCCGTGTGAGTGCACCAAGAGAAACAGTCAACAGT	763
QY	825	GAAGCACTGTTTAAACCAAGAGGTTTCAGACCCCTGTGCCCGGTTTTTGGAGGAGCCGAAGG	884
Db	764	AAAGGAAGAGCTGAGCCAGAAAGTTTGAACAGTTGTATCCATCTCGTGATGAGAAAGAAAG	823
QY	885	CGAAGTGGTTTAAGCACCTTGGCGCCGGGAGCCAGAGAGAGAAAGTTTGCAGCGCGTGGCGGCT	944
Db	824	TGAGTGTCTGCCAGGGAATCAGCAGAGGACGGAGAAAAGTTAGTTC--ATCAGAGGCT	880
QY	945	CATCCGCACTAGCG	959
Db	881	CATCCAGCATACAG	895

RESULT 14	
BM726338	
LOCUS	658 bp
DEFINITION	UI-E-EJ0- <i>alh-n-03-0-UI.r1</i> UI-E-EJ0 Homo sapiens linear EST 01-MAR-2002
ACCESSION	UI-E-EJ0- <i>alh-n-03-0-UI</i> 5', mRNA sequence.
VERSION	BM726338
KEYWORDS	BM726338.1 GI:19047671
SOURCE	EST.
ORGANISM	human.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE	1 (bases 1 to 658)
AUTHORS	Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery



JOURNAL  
MEDLINE  
COMMENT

Genome Res. 6 (9), 791-806 (1996)  
9704477  
Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@iue.uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
Seq primer: M13 Reverse.

FEATURES  
source

1. 658  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJ0-ah-n-03-0-UI"  
/clone\_lib="UI-E-EJ0"  
/tissue\_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/note="Organ: eye; Vector: pPT73-Pac (Pharmacia) with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I; UI-E-EJ0 is a subcloned cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pPT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTGT; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 177 a 166 c 196 g 116 t 3 others

ORIGIN

Query Match 24.1%; Score 345.4; DB 14; Length 658;  
Best Local Similarity 86.2%; Pred. No. 3.5e-65;  
Matches 405; Conservative 0; Mismatches 59; Indels 6; Gaps 2;

QY 118 AACTGACCGGGAAGGTGACAGGCGAG-GGAGAGAGGCCAAGCAGAGCGAG 176  
DB 194 ATCTAAGCGGGAAGGTCTACAGGCGAGGTGAGGCGAGGAGCGCCAGGCGAG 253  
177 GCAGGCTCCAGAGCGCGCGGATGACGCTTACGCGTTCAGCGCTCTAGGGGA 236  
DB 254 GCACGAGC-----ACCGAGGGGATGAGACTTCAGAGTGGTTTCAAGCGCGTCTAGGGGA 308  
QY 237 TGGCGACACATGACACACTTTGGAGAGCAGCTATTGGCCCATCTGCTGGAGATGTT 296  
DB 309 TGCACACAGCATGACACACTTCGAGAGACACTCATCTGCCCCATCTCCGTGGAATGTT 368  
QY 297 CTCACAGCCCGTGTGATCTTGGCTGCGACACACACTTGGCGCAGTGGCAACGA 356  
DB 369 CTCACACAGCAGTGTGATCTGCTGCGCAGCAGACACTTGGCCCAATATGGCAACGA 428  
QY 357 CGTCTTCAGAGCCTCTAATCTCTGTGGCAATCCCGGGGCTCCACACAGTGTCTTCAGG 416  
DB 429 CGTCTTCAGAGCCTCTAATCTCTGTATGAGTCTGCGGAGTCTCCACACTGTGTCTTCAGG 488  
QY 417 AGGACGTTTCCGATGCCCATCTTGTAGGACAGAGTGTCTCTGACACAGGAGTGTCTGA 476

DB 489 AGCCGTTTCCGCTGCCACATGTCAGCAGCATAGAGTGTCTGACAGACAGGTCCTGA 548  
QY 477 TGGCCTGACGGGAACCTGTCAGTGGAGACATCATTTACATCTACAGCAGAGTCTC 536  
DB 549 CGGCTGAGGGAACCTGTCAGTGGAGACATTTATCGACATTTACACAGAGTCTATC 608  
QY 537 CCGGCCACTGACCGCCAGGCTGGAACAGCAGCTCATGTGTGTGAGGAGC 586  
DB 609 CAGCCGCTGCACTCCAGCGCTGACAGCAGCACTCATGTGTGAGGAGCATG 658

RESULT 15  
BI754358 793 bp mRNA linear EST 25-SEP-2001  
BI754358  
LOCUS 603026906P1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5197560 5',  
DEFINITION mRNA sequence.  
ACCESSION BI754358  
VERSION BI754358 GI:15745936  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 793)  
NIH-MGC http://mgc.ncl.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L14M11494 row: k column: 01  
High quality sequence stop: 793.

FEATURES  
source

1. 793  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5197560"  
/clone\_lib="NIH\_MGC\_114"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-Sport6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH\_MGC Library."

BASE COUNT 201 a 208 c 228 g 156 t

ORIGIN

Query Match 23.9%; Score 342.6; DB 13; Length 793;  
Best Local Similarity 72.1%; Pred. No. 1.6e-64;  
Matches 474; Conservative 0; Mismatches 179; Indels 4; Gaps 2;

QY 226 CTGCTAGGGATGCGCAACATGAGACACTTGGAGAGCACTTGGCCCATGCGC 285  
DB 140 CTGATCCAGATGAGGATCCATGAGAACTTGGAGAGCAGCTGATCTGCTATCTGC 199  
QY 286 CTGAGATGTTTCCAAAGCCGCTGATCTTGGCTTCCCAACAACTGTGCGGCAAG 345  
DB 200 CTGAGATGTTTACCAAGCAGTGTGATCTTGGCGGCGACGACAACTGTGCGGAA 259  
QY 346 TGTGCAACGAGCTCTTCAAGGCTCTAATCTCTGTGGCAATCCGGGGCTCCACAAG 405  
DB 260 TGTGCAATGACATCTTCCAGGCTGCAAAATCCCTACTGTGACCAAGCGGGGAGCTCA 319

QY	406	GTGTCCTCAGGAGAGAGCTTTCGATCCCATCTCTTGAAGCAGAGGTTGCTCGACAG	465
Dp	320	TTCCATGTCTGGAGAGCCGTTTCCGCTGCCCCACCCTGCCGACGAGGTGATCATGATTCGT	379
QY	466	CATGCTGTCTATGAGCCTCGACGGGAACTGCTAGTGGAGAACATCATTCATCTCAAG	525
Dp	380	CACGAGAGTGTACGGCCTCGACAGAGAACTGTGTGGAGAACATCATTCATCTCAAG	439
QY	526	CAGGAGTCTCTCCCGGCGCATTCGACCGCCAAAGCGTGAACACGACCTCATCTGTGTAGAGAC	585
Dp	440	CAGGAGTCTCTCCAGTGGCGGCGCTCGACAAAGGGCA -- GTACCCCCATTCGCAAGGAGCAC	496
QY	586	GAGGACGAGAAAGATTAACATCTACTGCTCTGAGACTGAGAGTGCCGCCACCTGCTCTCTG	645
Dp	497	GAAGATGAGAAATCAACATCTACTGTCTCAGGTGTGAGAGTGCCGCCACCTGCTCTCTCTG	556
QY	646	AAGGTTTTCGGCGCCCAACAAGACCTGTGAGGTGGCCCCCTGCCCCACATTTACAACGC	705
Dp	557	AAGGTTTTCGGGATCCACAAGGCTTCGCGAGGTGGCCCCCATTCAGAGTGTCTTCAGAGGA	616
QY	706	CAGAGAGTGTGAGCTGAGCGATGATGCAATCGGAGATGTGGTGCGGGCAATATACCGTGTGAG	765
Dp	617	CAAAAGACTGAAACTAATAATCTATCTTCATCTGCTGTGTGGCGGGGAATATACCGTGTGAG	676
QY	766	GCATGTATCACCAGATGAGAGGAGGTGTGCCAGACCAATTGAGAGACACAG - CCGAGACA	824
Dp	677	AACATCATCTACCTACACTGGAGAGATTCCCGTCAGATGACCAAGGAACAACAGTACACAGGTA	736
QY	825	GAAGCAATCTGTTAAACACAGAGGTTGAGAGCCCTGAGCGGTTTGGAGAGACCGAA	881
Dp	737	CACGGAAGAGCTGAGCCCAAGAACTCTGACACAGTGTATGACCATCTCGAGATGAGAAAGAA	793

REFERENCE	TITLE	JOURNAL	COMMENT
AW918753			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
<p>AW918753 378 bp mRNA linear EST 25-MAY-2000</p> <p>EST350057 Rat gene index, normalized rat, norvegicus, Bento Soares</p> <p>Rattus norvegicus cDNA clone RG1RX57 5' end, mRNA sequence.</p> <p>AW918753</p> <p>AW918753.1 GI:8084538</p> <p>EST.</p> <p>Norway rat.</p> <p>Rattus norvegicus</p> <p>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;</p> <p>Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;</p> <p>Rattus.</p> <p>1 (bases 1 to 378)</p> <p>Lee,N.H., Glodok,A., Chandra,I., Mason,T.M., Quackenbush,J.,</p> <p>Kerlavage,A.R. and Adams,M.D.</p> <p>Rat Genome Project: Generation of a Rat EST (RESt) Catalog &amp; Rat</p> <p>Gene Index</p> <p>Unpublished (1998)</p> <p>Contact: Lee, NH</p> <p>The Institute for Genomic Research</p> <p>9712, Medical Center Drive, Rockville, MD 20850, USA</p> <p>Tel: (301)-838-3528</p> <p>Fax: (301)-838-0208</p> <p>Email: nhlee@tigr.org</p> <p>This clone is available through the ATCC</p> <p>tel#703-365-2700 for further information</p> <p>Seq primer: M13 Reverse.</p> <p>Location/Qualifiers</p> <p>1..378</p>			

```

/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="RG1EX57"
/clone_lib="Rat gene index, normalized rat, norvegicus,
Bento Soares"
/tissue_type="mix - brain, ovary, placenta, kidney, lung
liver, embryo, heart, muscle, spleen"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Estimated insert size approx.1 kb"

```

BASE COUNT	87 a	100 c	122 g	69 t	
Query Match		23.9%;	Score 341.8;	DB 10;	Length 378;
Best Local Similarity		94.2%;	Pred. No. 1.7e-64;		
Matches 355;	Conservative	0;	Mismatches 22;	Indels	0;
				Gaps	0;
QY	599	TCACATCTACTACGCTCTGAGCTGCGAGAGTCCACACCGCTCTCTCTGCAAGGTTTTCGGCG	658		
Db	2	TTAATCATCTACTGTCTGAGCTGTGAGAGTCCACACCGCTCTCTCTGCAAGCTTTTGGCG	61		
QY	659	CCCAACAGSACTGTAGGTGGGCGCCCTCTGCGCCACCATTTACAAGCCGAGAGAGTGAC	718		
Db	62	CCCAACAGSACTGTAGGTGGGCGCCCTCTGCGCCACCATTTACAAGCCGAGAGAGTGAC	121		
QY	719	TGAGCGATGCGATCGCGATGCTGTTGGCGCGGCAATGACCGTGTGCAAGCAGTATCACCC	778		
Db	122	TGAGTGACGCGCTCCGCGATGCTGTTGGCGCGGCAATGACCGTGTGCAAGCAGTATCACCC	181		
QY	779	AGATGAGAGAGTGTGCCAGACCATTTGAGAGCAACAGCCGCGACAAGAACGACTGTTA	838		
Db	182	AGATGAGAGAGTGTGCCAGACCATTTGAGAGCAACAGCCGCGACAAGAACGACTGTTA	241		
QY	839	ACCAGAGGTTGAGACCCCTGTGCGGGTTTTTGGAGAGACGCAAGGCGCAACGTCTTCAAG	898		
Db	242	ACCAGAAATTCAGACCCCTGTGCGGGTTTTTGGAGAGACGCAAGGCGCAAGTGTCTTCAAG	301		
QY	899	CACCTGGCCCGGAGCAGAGAGTGTGCAGCCGGTGGGGGGCTCTATCCGCCCAAGTACG	958		
Db	302	CGCTGGCTCGGGTGCACAAAGGAGGAAGTGTGCAGCCGGTGTAGCTCATCCGCCAGTACG	361		
QY	959	GAGACCACTTGGAGGCG	975		
Db	362	GAGACCACTTGGAGGCG	378		

RESULT	17
LOCUS	BG375657/c
DEFINITION	BG375657 440 bp mRNA linear EST 12-MAR-2001
ACCESSION	U1-R-CS0-bld-h-03-0-U1.s1 U1-R-CS0 Rattus norvegicus cDNA clone
VERSION	U1-R-CS0-bld-h-03-0-U1.3 , mRNA sequence.
KEYWORDS	BG375657.BG375657.1 GI:13300129
SOURCE	EST.
ORGANISM	Norway rat. Rattus norvegicus
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
TITLE	1 (bases 1 to 440) Bonaldo,M.F., Lennon,G. and Soares,M.B.
JOURNAL	Normalization and subtraction: two approaches to facilitate gene discovery
MEDLINE	Genome Res. 6 (9), 791-806 (1996)
COMMENT	9704477 Contact: Soares, MB

Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9365  
Email: mscares@blue.weeg.uiowa.edu  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. The sequence tag present in the cDNA between the NotI site  
and the oligo-dT track served to verify it as a clone from the  
normalized rat heart pool library cDNA library Preparation: M. B.  
Soares Lab Clone distribution: clones will be available through  
Research Genetics ([www.resgen.com](http://www.resgen.com))  
Seq primer: M13 Forward  
POLY=Yes  
Location/Qualifiers

## FEATURES

Location/Qualifiers

source 1. .440  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-CS0-btd-b-03-0-UI"  
 /clone\_11b="UI-R-CS0"  
 /dev\_stage="ADULT"  
 /lab\_host="DHI0B (Life Technologies)"  
 /note="Vector: p7T3D-Pac (Pharmacia) with a modified polylinker. Site\_1: Not I; Site\_2: Eco RI; The UI-R-CS0 library is a normalized library constructed from the following rat heart tissues: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, adult day 200. For a detailed description of the library from which this clone was derived, please visit our web site at [rattest.eng.uiowa.edu](http://rattest.eng.uiowa.edu). The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
 TAG\_LIB=UI-R-CS0  
 TAG\_TISSUE=rat heart pool  
 TAG\_SEQ=ATAGATAC"

BASE COUNT 69 a 144 c 113 g 114 t  
 ORIGIN

Query Match 23.8%; Score 341.4; DB 12; Length 440;  
 Best Local Similarity 90.3%; Pred. No. 2.2e-64;  
 Matches 391; Conservative 0; Mismatches 31; Indels 11; Gaps 2;

QY 999 CATCCAGTCCATGAGAGAGCCGCGATGCTCTCTACCTCCAGCAGCAAGAGCTGAT 1058  
 |||||||  
 DB 440 CATCCAGTCCATGAGAGAGCCGCGATGCTCTCTACCTCCAGCAGCAGAGCTGAT 381  
 |||||||

QY 1059 CAACAAGTGGGGGCGCATGTCGAAAGTGGAGCTGGAGAGAGCCGCGAGCGGCTATGA 1118  
 |||||||  
 DB 380 CAACAAGTGGGGGCGCATGTCGAAAGTGGAGCTGGAGAGCCGCGAGCGGCTATGA 321  
 |||||||

QY 1119 GAGCATGAGCAATTCCTGTGTGAGCTGGAGCAGTGGCCGAATGTTGGCAACCATCGA 1178  
 |||||||  
 DB 320 GAGCATGAGCAATTCCTGTGTGAGCTGGAGCAGTGGCCGAATGTTGGCGACATCGA 261  
 |||||||

QY 1179 CTTCGAGCCGGGCGCGCTGGGGATGAAGATGACGACATGCTTTGGATGGGAGGA 1238  
 |||||||  
 DB 260 CTTCGAGCCGGGCGCGCTGGGGATGAAGATGACGACATGCTTTGGAGCGGAGGA 201  
 |||||||

QY 1239 GGGCATGGGGGCTGGAGAGAGGGGCTGGAGCTGCCAAGAGGCTCAGGCTGACCTG 1298  
 |||||||  
 DB 200 GGGCAACACAGGGCTGGAGAGAGGGGCTGGAGCGGCGCAAGAGGTT-----TGCATG 147  
 |||||||

QY 1299 ACCGCACTGATCCAGAGCGCACACCCGAAGCGGAGCAAGGATGCTGAGGATCGC 1358  
 |||||||  
 DB 146 ATCCGACTGATCCAGAGCGCACGCC-----CGGAGCCCAAGGATGCTGAGGATCGC 92  
 |||||||

QY 1359 GCAGAGACCAACCGCGCCACCAAGCTGGCTTCCGCCCCCGGGAAGGTTCTCAATAAAG 1418  
 |||||||  
 DB 91 GCAGAGACCAACCGCGCTACCCATCTCGCTCCACCCCGAGGATGTTCTCAATAAAGA 32  
 |||||||

QY 1419 ACTCAGTGTCCC 1431  
 |||||||  
 DB 31 ACTCAGTGTGTCCC 19  
 |||||||

RESULT 18  
 BG379943/c 441 bp mRNA linear EST 12-MAR-2001  
 LOCUS BG379943  
 DEFINITION UI-R-CS0-btd-b-01-0-UI.s1 UI-R-CS0 Rattus norvegicus cDNA clone.  
 ACCESSION BG379943  
 VERSION BG379943.1 GI:13304415  
 KEYWORDS EST.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine;

REFERENCE 1 (bases 1 to 441)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: [masores@blue.wee.uiowa.edu](mailto:masores@blue.wee.uiowa.edu)  
 The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA library preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics ([www.resgen.com](http://www.resgen.com))  
 Seq primer: M13 Forward  
 POLYA=yes.

FEATURES  
 Location/Qualifiers  
 source 1. .441  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley"  
 /db\_xref="taxon:10116"  
 /clone="UI-R-CS0-btd-b-01-0-UI"  
 /clone\_11b="UI-R-CS0"  
 /dev\_stage="ADULT"  
 /lab\_host="DHI0B (Life Technologies)"  
 /note="Vector: p7T3D-Pac (Pharmacia) with a modified polylinker. Site\_1: Not I; Site\_2: Eco RI; The UI-R-CS0 library is a normalized library constructed from the following rat heart tissues: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, adult day 200. For a detailed description of the library from which this clone was derived, please visit our web site at [rattest.eng.uiowa.edu](http://rattest.eng.uiowa.edu). The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)  
 TAG\_SEQ=None found"

BASE COUNT 69 a 148 c 110 g 114 t  
 ORIGIN

Query Match 23.8%; Score 340.6; DB 12; Length 441;  
 Best Local Similarity 90.7%; Pred. No. 3.3e-64;  
 Matches 389; Conservative 0; Mismatches 29; Indels 11; Gaps 2;

QY 1003 CATCCATGAGAGAGCCGCGATGCTCTCTACCTCCAGCAGCAGCAAGAGCTGATCAAC 1062  
 |||||||  
 DB 441 CATCCATGAGAGAGCCGCGATGCTCTCTACCTCCAGCAGCAGCAAGAGCTGATCAAC 382  
 |||||||

QY 1063 AAGTCTGGGGCAATGTCGAAGTGGAGCTGGAGAGCCGCGAGCGGCTATGAAGC 1122  
 |||||||  
 DB 381 AAGTCTGGGGCAATGTCGAAGTGGAGCTGGAGAGCGCGGAGCGGCTATGAAGC 322  
 |||||||

QY 1123 ATGAGCAATTCCTGTGTGAGCTGGAGCAGTGGCGGAATGTTGGCAACCATGACTTC 1182  
 |||||||  
 DB 321 ATGAGCAATTCCTGTGTGAGCTGGAGCAGTGGCGGAATGTTGGCGACATGACTTC 262  
 |||||||

QY 1183 CAGCCGGGCGCGCTGGGGATGAAGAGATGACGACATGCTTTGGATGGGAGAGGCG 1242  
 |||||||  
 DB 261 CAGCCGGGCGCGCTGGGGATGAAGAGATGACGACATGCTTTGGAGCGGGAAGGCG 202  
 |||||||

QY 1243 AATGGCGGCTGGAGAGAGCGGCTGGAGCTGCCAAGAGCTCAGGCTGCTGAGACCC 1302  
 |||||||  
 DB 201 AACACAGGCTGGAGAGAGCGGCTGGAGCGGCGCAAGAGTT-----TGCATGATCC 148  
 |||||||

QY 1303 GACTTGATCCAGAGCGCACACCCGAAGCGGAGCCCAAGGATGCTGAGATCTGCGCAG 1362  
 |||||||  
 DB 147 GACTTGATCCAGAGCGCACGCC-----CGGAGCCCAAGGATGCTGAGATCTGCGCAG 93  
 |||||||

QY 1363 AGACCAACCGCGCCACCAAGCTCGGCTTCCGCCCCCGGGAAGGTTCTCAATAAAGACTC 1422  
 |||||||







GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 09:13:42 ; Search time 59 Seconds

(without alignments) updates/sec  
7438.211 Million cell

Title: US-09-908-988B-1

Perfect score: 1431  
Sequence: 1 aagagagtgtacagagagtgt.....ataaagactcaagtgtccc 1431Scoring table: OLIGO-MUC  
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size : 15

Total number of hits satisfying chosen parameters: 744

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents, NA:\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCRTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/Backfillseq1.seq:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	1.2	39	3 US-08-767-942A-36	Sequence 36, Appl
2	17	1.2	102	1 US-08-047-041A-15	Sequence 15, Appl
3	17	1.2	133	1 US-08-047-041A-2	Sequence 2, Appl
4	17	1.2	133	2 US-08-795-006A-22	Sequence 22, Appl
5	17	1.2	133	4 US-09-184-073-22	Sequence 22, Appl
6	17	1.2	320	3 US-09-157-177-116	Sequence 116, App
7	17	1.2	716	4 US-08-998-416-1129	Sequence 1129, App
8	17	1.2	732	4 US-08-998-416-927	Sequence 927, App
9	17	1.2	800	1 US-08-472-239-1	Sequence 1, Appl
10	17	1.2	1070	4 US-09-414-436-4	Sequence 4, Appl
11	17	1.2	1215	1 US-08-347-792-20	Sequence 20, Appl
12	17	1.2	1215	1 US-08-431-357-20	Sequence 20, Appl
13	17	1.2	1215	1 US-08-392-542-26	Sequence 26, Appl
14	17	1.2	1215	4 US-08-894-327-26	Sequence 26, Appl
15	17	1.2	1215	4 US-09-685-027-26	Sequence 26, Appl
16	17	1.2	1215	5 PCT-US95-15353-20	Sequence 20, Appl
17	17	1.2	1303	1 US-08-047-041A-13	Sequence 13, Appl
18	17	1.2	1307	1 US-08-047-041A-12	Sequence 12, Appl
19	17	1.2	1317	1 US-07-912-011-1	Sequence 1, Appl
20	17	1.2	1317	1 US-08-347-792-1	Sequence 1, Appl
21	17	1.2	1317	1 US-08-431-357-1	Sequence 1, Appl
22	17	1.2	1317	2 US-08-697-221-1	Sequence 1, Appl
23	17	1.2	1317	4 US-08-392-542-1	Sequence 1, Appl
24	17	1.2	1317	4 US-08-894-327-1	Sequence 1, Appl
25	17	1.2	1317	4 US-09-147-751-8	Sequence 8, Appl
26	17	1.2	1317	4 US-09-305-914-1	Sequence 1, Appl
27	17	1.2	1317	4 US-09-685-027-1	Sequence 1, Appl

28	17	1.2	1317	5 PCT-US95-15353-1	Sequence 1, Appl
29	17	1.2	1683	1 US-07-945-283-3	Sequence 3, Appl
30	17	1.2	1710	1 US-07-903-103-3	Sequence 3, Appl
31	17	1.2	1710	1 US-08-044-619A-3	Sequence 3, Appl
32	17	1.2	1710	1 US-08-283-911-3	Sequence 3, Appl
33	17	1.2	1710	1 US-08-245-500A-4	Sequence 4, Appl
34	17	1.2	1710	1 US-08-390-546-4	Sequence 4, Appl
35	17	1.2	1710	1 US-08-390-479A-4	Sequence 4, Appl
36	17	1.2	1710	1 US-08-557-393-4	Sequence 4, Appl
37	17	1.2	1710	1 US-08-390-516C-4	Sequence 4, Appl
38	17	1.2	1710	1 US-08-390-517A-4	Sequence 4, Appl
39	17	1.2	1710	1 US-08-390-515A-4	Sequence 4, Appl
40	17	1.2	1710	2 US-08-801-718-4	Sequence 4, Appl
41	17	1.2	1710	3 US-09-073-567-12	Sequence 12, Appl
42	17	1.2	1710	4 US-09-170-159A-4	Sequence 4, Appl
43	17	1.2	1710	4 US-09-480-718-45	Sequence 45, Appl
44	17	1.2	1722	4 US-09-385-028-15	Sequence 15, Appl
45	17	1.2	1757	3 US-08-094-071-1	Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-08-767-942A-36  
Sequence 36, Application US/08767942A  
Patent No. 6068982

GENERAL INFORMATION:  
APPLICANT: Rolfe, Mark  
APPLICANT: Chin, M. Isabel  
APPLICANT: Berlin, Vivian  
APPLICANT: Damagnez, Veronique  
APPLICANT: Draetta, Giulio  
APPLICANT: Guillaume, Cottarel  
TITLE OF INVENTION: UNROUTIN COMUGATING ENZYMES  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/767, 942A  
FILING DATE: 17-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-029.04  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-7000  
TELEFAX: 617-832-1000

INFORMATION FOR SEQ. ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 39 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "primer"

US-08-767-942A-36

Query Match 1.28; Score 17; DB 3; Length 39;  
Best local Similarity 100.0%; Pred. No. 87;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1007 CCATGGAGGAGCCGACG 1023

DB 14 CCATGAGAGAGCCGCGAG 30

|||||

## RESULT 2

US-08-047-041A-15

Sequence 15, Application US/08047041A

Patent No. 5527676

GENERAL INFORMATION:

APPLICANT: Vogelstein, Bert

APPLICANT: Baker, Suzanne J.

APPLICANT: Fearon, Eric R.

APPLICANT: Nigro, Janice M.

TITLE OF INVENTION: Detection of Loss of the Wild-Type p53

TITLE OF INVENTION: Gene

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Banner & Allegretti, Ltd.

STREET: 1001 G Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20001-4597

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/047,041A

FILING DATE: 22-MAR-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/928,661

FILING DATE: 17-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/446,584

FILING DATE: 06-DEC-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/330,566

FILING DATE: 29-MAR-1989

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.

REGISTRATION NUMBER: 32,141

REFERENCE/DOCKET NUMBER: 01107,42917

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-508-9100

TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 102 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

POSITION IN GENOME:

CHROMOSOME/SEGMENT: exon 2

PUBLICATION INFORMATION:

AUTHORS: Lamb,

JOURNAL: Mol. Cell. Biol.

VOLUME: 6

ISSUE: 5

PAGES: 1379-1385

DATE: 1986

US-08-047-041A-15

Query Match

Best Local Similarity 1.2%; Score 17; DB 1; Length 102;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGAGAGAGCCGCGAG 1023

|||||

DB 27 CCATGAGAGAGCCGCGAG 43

## RESULT 3

US-08-047-041A-2

Sequence 2, Application US/08047041A

Patent No. 5527676

GENERAL INFORMATION:

APPLICANT: Vogelstein, Bert

APPLICANT: Baker, Suzanne J.

APPLICANT: Fearon, Eric R.

APPLICANT: Nigro, Janice M.

TITLE OF INVENTION: Detection of Loss of the Wild-Type p53

TITLE OF INVENTION: Gene

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Banner & Allegretti, Ltd.

STREET: 1001 G Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20001-4597

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/047,041A

FILING DATE: 22-MAR-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/928,661

FILING DATE: 17-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/446,584

FILING DATE: 06-DEC-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/330,566

FILING DATE: 29-MAR-1989

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.

REGISTRATION NUMBER: 32,141

REFERENCE/DOCKET NUMBER: 01107,42917

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-508-9100

TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 133 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

POSITION IN GENOME:

CHROMOSOME/SEGMENT: exon 2

PUBLICATION INFORMATION:

AUTHORS: Buchman, V. L.

TITLE: A variation in the structure of the

JOURNAL: Gene

VOLUME: 70

PAGES: 245-252

DATE: 1988

US-08-047-041A-2

Query Match

Best Local Similarity 1.2%; Score 17; DB 1; Length 133;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGGAGAGCCGCGAG 1023  
|||||  
Db 42 CCATGGAGAGCCGCGAG 58

## RESULT 4

US-08-795-006A-22  
; Sequence 22, Application US/08795006A  
; Patent No. 5840579  
; GENERAL INFORMATION:  
; APPLICANT: Boeke, Jef  
; APPLICANT: Brachmann, Rainer  
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING P53  
; TITLE OF INVENTION: MUTATIONS WHICH SUPPRESS P53 CANCER MUTATIONS  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff  
; STREET: 1001 G Street, NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/795,006A  
; FILING DATE: 05-FEB-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A  
; REGISTRATION NUMBER: 32141  
; REFERENCE/DOCKET NUMBER: 01107.03170  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 133 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-795-006A-22

Query Match 1.2%; Score 17; DB 2; Length 133;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGGAGAGCCGCGAG 1023  
|||||  
Db 42 CCATGGAGAGCCGCGAG 58

## RESULT 5

US-09-184-073-22  
; Sequence 22, Application US/09184073  
; Patent No. 6183964  
; GENERAL INFORMATION:  
; APPLICANT: Boeke, Jef  
; APPLICANT: Brachmann, Rainer  
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING P53  
; TITLE OF INVENTION: MUTATIONS WHICH SUPPRESS P53 CANCER MUTATIONS  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner & Witcoff

STREET: 1001 G Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/184,073  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/795,006  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A  
; REGISTRATION NUMBER: 32141  
; REFERENCE/DOCKET NUMBER: 01107.03170  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 133 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-184-073-22

Query Match 1.2%; Score 17; DB 4; Length 133;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGGAGAGCCGCGAG 1023  
|||||  
Db 42 CCATGGAGAGCCGCGAG 58

## RESULT 6

US-09-157-177-116  
; Sequence 116, Application US/09157177  
; Patent No. 6090558  
; GENERAL INFORMATION:  
; APPLICANT: Butler, John M.  
; APPLICANT: Li, Jia  
; APPLICANT: Monforte, Joseph A.  
; APPLICANT: Becker, Christopher H.  
; TITLE OF INVENTION: DNA TYPING BY MASS SPECTROMETRY WITH POLYMORPHIC DNA  
; FILE REFERENCE: GETR:017/GER017P  
; CURRENT APPLICATION NUMBER: US/09/157,177  
; CURRENT FILING DATE: 1998-09-18  
; EARLIER APPLICATION NUMBER: 60/059,415  
; EARLIER FILING DATE: 1997-09-19  
; NUMBER OF SEQ ID NOS: 135  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 116  
; LENGTH: 320  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-157-177-116

Query Match 1.2%; Score 17; DB 3; Length 320;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 ATAGACAGGGGTGAGAG 43  
|||||  
Db 237 ATAGACAGGGGTGAGAG 253

RESULT 7  
US-08-998-416-1129  
; Sequence 1129, Application US/08998416  
; Patent No. 6239264  
; GENERAL INFORMATION:  
; APPLICANT: Philippsen, Peter  
; APPLICANT: Pohlmann, Rainer  
; APPLICANT: Steiner, Sabine  
; APPLICANT: Mohr, Christine  
; APPLICANT: Wendland, Jurgen  
; APPLICANT: Knechtle, Philipp  
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPIT  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 1152  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 6239264rtis Corporation  
; STREET: 3054 Cornwalis Road  
; CITY: Research Triangle Park  
; STATE: No. 6239264th Carolina  
; COUNTRY: USA  
; ZIP: 27709  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/998,416  
; FILING DATE: 24-DEC-1997  
; CLASSIFICATION: 435  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: CH 0016/97  
; FILING DATE: 31-DEC-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meigs, J. Timothy  
; REGISTRATION NUMBER: 38,241  
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-541-8587  
; TELEFAX: 919-541-8689  
; INFORMATION FOR SEQ ID NO: 1129:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 716 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: PAG1688RP  
; US-08-998-416-1129

Query Match 1.2%; Score 17; DB 4; Length 716;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1192 GCCGCTGGGGATGAGA 1208  
|||||

DB 526 GCCGCTGGGGATGAGA 542

RESULT 8  
US-08-998-416-927  
; Sequence 927, Application US/08998416  
; Patent No. 6239264  
; GENERAL INFORMATION:  
; APPLICANT: Philippsen, Peter  
; APPLICANT: Pohlmann, Rainer  
; APPLICANT: Steiner, Sabine  
; APPLICANT: Mohr, Christine  
; APPLICANT: Wendland, Jurgen

APPLICANT: Knechtle, Philipp  
APPLICANT: Reibischung, Corinne  
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPIT  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 1152  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6239264rtis Corporation  
STREET: 3054 Cornwalis Road  
CITY: Research Triangle Park  
STATE: No. 6239264th Carolina  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/998,416  
FILING DATE: 24-DEC-1997  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: CH 0016/97  
FILING DATE: 31-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 927:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 732 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: PAG1576RP  
US-08-998-416-927

Query Match 1.2%; Score 17; DB 4; Length 732;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1192 GCCGCTGGGGATGAGA 1208  
|||||

DB 526 GCCGCTGGGGATGAGA 542

RESULT 9  
US-08-472-239-1  
; Sequence 1, Application US/08472239  
; Patent No. 5728526  
; GENERAL INFORMATION:  
; APPLICANT: GEORGE, Jr., Albert L.  
; APPLICANT: BHATNAGAR, Satish K.  
; APPLICANT: NAZARENKO, Irena  
; TITLE OF INVENTION: METHOD FOR ANALYZING A NUCLEOTIDE  
; TITLE OF INVENTION: SEQUENCE AND KIT THEREFOR  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472.239  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mukai, Robert G.  
REGISTRATION NUMBER: 28,531  
REFERENCE/DOCKET NUMBER: 020160-229  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 800 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-472-239-1

Query Match 1.2%; Score 17; DB 1; Length 800;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGGAGAGCCGCAG 1023  
|||||  
DB 115 CCATGGAGAGCCGCAG 131

RESULT 10  
US-09-414-436-4  
Sequence 4, Application US/09414436  
Patent No. 6294384  
GENERAL INFORMATION:  
APPLICANT: Dell'Acqua, Giorgio  
APPLICANT: Mann, Michael J.  
APPLICANT: Drau, Victor J.  
TITLE OF INVENTION: Compositions and Methods Based Upon an Isoform of p53  
FILE REFERENCE: p53f  
CURRENT APPLICATION NUMBER: US/09/414,436  
CURRENT FILING DATE: 1999-10-07  
EARLIER APPLICATION NUMBER: 60/103,849  
EARLIER FILING DATE: 1999-10-09  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 4  
LENGTH: 1070  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-414-436-4

Query Match 1.2%; Score 17; DB 4; Length 1070;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGGAGAGCCGCAG 1023  
|||||  
DB 134 CCATGGAGAGCCGCAG 150

RESULT 11  
US-08-347-792-20  
Sequence 20, Application US/08347792  
Patent No. 5573925  
GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos D.  
TITLE OF INVENTION: p53 Proteins with Altered  
Tetramerization Domains  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr., PO Box 457

CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/347,792  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST580USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1215 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-347-792-20

Query Match 1.2%; Score 17; DB 1; Length 1215;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGGAGAGCCGCAG 1023  
|||||  
DB 26 CCATGGAGAGCCGCAG 42

RESULT 12  
US-08-431-357-20  
Sequence 20, Application US/08431357  
Patent No. 5721340  
GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos D.  
TITLE OF INVENTION: p53 Proteins with Altered  
Tetramerization Domains  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr., PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,357  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/347,792  
FILING DATE: 28-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST580USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1215 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-431-357-20

Query Match 1.2%: Score 17; DB 1; Length 1215;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGAGAGAGCCGCAG 1023  
|||||  
DB 26 CCATGAGAGAGCCGCAG 42

RESULT 13  
US-08-392-542-26

Sequence 26, Application US/08392542  
Patent No. 6169073  
GENERAL INFORMATION:

APPLICANT: Halazonetis, Thanos  
APPLICANT: Hartwig, Wolfgang  
TITLE OF INVENTION: Peptides and Peptidomimetics with  
STRUCTURAL SIMILARITY TO HUMAN P53 THAT ACTIVATE P53  
TITLE OF INVENTION: Function  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie & Beckett  
STREET: 1001 G Street, N.W.  
CITY: Washington, D.C.  
STATE: District of Columbia  
COUNTRY: U.S.  
ZIP: 20001

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/392,542  
FILING DATE:

CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Posorske, Laurence H.  
REGISTRATION NUMBER: 34,698  
REFERENCE/DOCKET NUMBER: 0486,48439  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 508-9100  
TELEFAX: 202 508-9299

INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1215 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-08-392-542-26

Query Match 1.2%: Score 17; DB 4; Length 1215;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGAGAGAGCCGCAG 1023  
|||||  
DB 26 CCATGAGAGAGCCGCAG 42

RESULT 14

US-08-894-327-26  
Sequence 26, Application US/08894327  
Patent No. 6245886  
GENERAL INFORMATION:

APPLICANT: Halazonetis, Thanos  
APPLICANT: Hartwig, Wolfgang  
TITLE OF INVENTION: Peptides and Peptidomimetics with  
STRUCTURAL SIMILARITY TO HUMAN P53 THAT ACTIVATE P53  
TITLE OF INVENTION: Function  
FILE REFERENCE: 2973, 19998  
CURRENT APPLICATION NUMBER: US/08/894,327  
CURRENT FILING DATE: 1997-12-04  
EARLIER APPLICATION NUMBER: pctus96/01535  
EARLIER FILING DATE: 1996-02-16  
EARLIER APPLICATION NUMBER: 08/392,542  
EARLIER FILING DATE: 1995-02-16  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 26  
LENGTH: 1215  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-08-894-327-26

Query Match 1.2%: Score 17; DB 4; Length 1215;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGAGAGAGCCGCAG 1023  
|||||  
DB 26 CCATGAGAGAGCCGCAG 42

RESULT 15  
US-09-685-027-26

Sequence 26, Application US/09685027  
Patent No. 6420118  
GENERAL INFORMATION:

APPLICANT: Halazonetis, Thanos  
APPLICANT: Hartwig, Wolfgang  
TITLE OF INVENTION: Peptides and Peptidomimetics with  
STRUCTURAL SIMILARITY TO HUMAN P53 THAT ACTIVATE P53  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie & Beckett  
STREET: 1001 G Street, N.W.  
CITY: Washington, D.C.  
STATE: District of Columbia  
COUNTRY: U.S.  
ZIP: 20001

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/685,027  
FILING DATE: 10-Oct-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/392,542  
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Posorske, Laurence H.  
REGISTRATION NUMBER: 34,698  
REFERENCE/DOCKET NUMBER: 0486,48439  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 508-9100  
TELEFAX: 202 508-9299

INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1215 base pairs

TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
SEQUENCE DESCRIPTION: SEQ ID NO: 26;  
US-09-685-027-26

Query Match 1.2%; Score 17; DB 4; Length 1215;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1007 CCATGAGAGAGCCGAG 1023  
Db 26 CCATGAGAGAGCCGAG 42

RESULT 16  
PCT-US95-15353-20

Sequence 20, Application PC/TUS9515353  
GENERAL INFORMATION:  
APPLICANT: The Wistar Institute of Anatomy  
APPLICANT: and Biology  
APPLICANT: Halazonetis, Thanos D.  
TITLE OF INVENTION: p53 proteins with altered  
TITLE OF INVENTION: Telomerization Domains  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr., PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/15353  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/347,792  
FILING DATE: 28-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/431,357  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/456,623  
FILING DATE: 01-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: MST58CPCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1215 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US95-15353-20

Query Match 1.2%; Score 17; DB 5; Length 1215;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1007 CCATGAGAGAGCCGAG 1023  
Db 26 CCATGAGAGAGCCGAG 42

RESULT 17  
US-08-047-041A-13  
Sequence 13, Application US/08047041A  
Patent No. 5527676

GENERAL INFORMATION:  
APPLICANT: Vogelstein, Bert  
APPLICANT: Baker, Suzanne J.  
APPLICANT: Fearon, Eric R.  
APPLICANT: Nigro, Janice M.  
TITLE OF INVENTION: Detection of loss of the wild-type p53  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Allegretti, Ltd.  
STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001.4597  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/047,041A  
FILING DATE: 22-MAR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/928,661  
FILING DATE: 17-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/446,584  
FILING DATE: 06-DEC-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/330,566  
FILING DATE: 29-MAR-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 01107.42917  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9299  
TELEFAX: 202-508-9100  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1303 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
POSITION IN GENOME:  
MAP POSITION: 17p13.1  
PUBLICATION INFORMATION:  
AUTHORS: Harris, N.  
JOURNAL: Mol. Cell. Biol  
VOLUME: 6  
ISSUE: 12  
PAGES: 4650-4656  
DATE: 1986

Query Match 1.2%; Score 17; DB 1; Length 1303;  
Best Local Similarity 100.0%; Pred. No. 81;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1007 CCATGAGAGCGCGCAG 1023

Db 120 CCATGAGAGCGCGCAG 136

## RESULT 18

US-08-047-041A-12

Sequence 12, Application US/08047041A

Patent No. 5527676

GENERAL INFORMATION:

APPLICANT: Vogelstein, Bert

APPLICANT: Baker, Suzanne J.

APPLICANT: Fearon, Eric R.

APPLICANT: Nigro, Janice M.

TITLE OF INVENTION: Detection of Loss of the Wild-Type p53

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Banner & Allegretti, Ltd.

STREET: 1001 G Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20001-4597

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/047,041A

FILING DATE: 22-MAR-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/928,661

FILING DATE: 17-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/446,584

FILING DATE: 06-DEC-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/330,566

FILING DATE: 29-MAR-1989

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.

REGISTRATION NUMBER: 32,141

REFERENCE/DOCKET NUMBER: 01107.42917

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-508-9100

TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1307 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

POSITION IN GENOME:

MAP POSITION: 17p13.1

PUBLICATION INFORMATION:

AUTHORS: Harris, N.

TITLE: Molecular basis for heterogeneity of the

JOURNAL: Mol. Cell. Biol.

VOLUME: 6

ISSUE: 12

PAGES: 4650-4656

DATE: 1986

US-08-047-041A-12

Query Match 1.2%; Score 17; DB 1; Length 1307;

Best Local Similarity 100.0%; Pred. No. 81;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1007 CCATGAGAGCGCGCAG 1023

Db 124 CCATGAGAGCGCGCAG 140

## RESULT 19

US-07-912-011-1

Sequence 1, Application US/07912011

Patent No. 5382510

GENERAL INFORMATION:

APPLICANT: Levine, Arnold J.

APPLICANT: Shenk, Thomas E.

APPLICANT: Finlay, Cathy A.

TITLE OF INVENTION: Probes for Detecting Mutant p53

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Imclone Systems Incorporated

STREET: 180 Varlick Street

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10014

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/912,011

FILING DATE: 10-JUL-1992

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 28,601

FILING DATE: 212-645-1405

TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1317 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-07-912-011-1

Query Match 1.2%; Score 17; DB 1; Length 1317;

Best Local Similarity 100.0%; Pred. No. 81;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1007 CCATGAGAGCGCGCAG 1023

Db 134 CCATGAGAGCGCGCAG 150

## RESULT 20

US-08-347-792-1

Sequence 1, Application US/08347792

Patent No. 5573925

GENERAL INFORMATION:

APPLICANT: Halazonetis, Thanos D.

TITLE OF INVENTION: p53 Proteins With Altered

TITLE OF INVENTION: Tetramerization Domains

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Cntr., PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/347,792  
APPLICATION NUMBER: US/08/347,792  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST580USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1317 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 136..1314  
US-08-347-792-1

Query Match 1.2%; Score 17; DB 1; Length 1317;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1007 CCATGAGAGAGCCGCAG 1023  
|||||  
DB 134 CCATGAGAGAGCCGCAG 150

Search completed: December 3, 2002, 10:42:46  
Job time : 94 secs







XX WPI: 2002-241506/29.  
 DR P-PSDB: ABB082375.  
 XX Novel muscle ring finger protein useful for drug screening, and for  
 PT diagnosing and treating diseases, particularly cardiomyopathies  
 XX Claim 4: Page 123-125; 134pp; English.

The sequence encodes murine muscle ring finger protein 1 (MURF-1). The  
 CC invention relates to a purified muscle ring finger (MURF) protein,  
 CC selected from MURF-1, MURF-2 and MURF-3. The MURF proteins of the  
 CC invention are involved in microtubule and intermediate filament  
 CC stabilisation of striated muscle cells and have cardiant activity.  
 CC The MURF proteins are useful for screening a candidate substance  
 CC for MURF protein-binding activity, in a cell, cell-free system or  
 CC in vivo, and its effect on interaction of MURF with microtubules,  
 CC of homodimerisation of MURF, MURF-1, MURF-2 or MURF-3 stabilisation  
 CC of microtubules, interaction of MURF with intermediate filaments,  
 CC e.g. desmin, vimentin and cytokeratin, and heterodimerisation of MURF.  
 CC The screened compounds are useful for treating and preventing  
 CC cardiac hypertrophy and heart diseases. MURF proteins are useful as  
 CC antigens to immunise animals for the production of antibodies.

SQ Sequence 1431 BP: 338 A: 384 C: 473 G: 236 T: 0 other:

Query Match 100.0%; Score 1431; DB 24; Length 1431;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AAGAGTGTATACAGAGTGTCTGAAATATAGACAGGGGTGAGAGAGCTTTAGGGGAAG 60  
 1 AAGAGTGTATACAGAGTGTCTGAAATATAGACAGGGGTGAGAGAGCTTTAGGGGAAG 60  
 61 GACAGGACTCTTCCAGAGGAGGAGCAATAGCCGGATCCCAAGATCCAGTCCCTAAAC 120  
 61 GACAGGACTCTTCCAGAGGAGGAGCAATAGCCGGATCCCAAGATCCAGTCCCTAAAC 120  
 121 TGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
 121 TGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
 181 GCTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240  
 181 GCTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240  
 241 GCTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
 241 GCTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
 301 AAGCCGCTGTATCTTCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
 301 AAGCCGCTGTATCTTCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
 361 TTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
 361 TTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
 421 CGTTTCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
 421 CGTTTCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
 481 CTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
 481 CTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
 541 CCACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
 541 CCACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
 601 AACATCTACTGCTGAGTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
 601 AACATCTACTGCTGAGTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660

QY 661 CACAAGACTGTAGAGTGGCCCTCTGCCCACCATTTACAAACGCCAAGAGTAGCTG 720  
 DB 661 CACAAGACTGTAGAGTGGCCCTCTGCCCACCATTTACAAACGCCAAGAGTAGCTG 720  
 QY 721 AGCGATGCGATGCGATGCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
 DB 721 AGCGATGCGATGCGATGCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
 QY 781 ATGAGAGGAGTGTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
 DB 781 ATGAGAGGAGTGTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
 QY 841 CAGAGGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
 DB 841 CAGAGGTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
 QY 901 CTGCGCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 DB 901 CTGCGCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
 QY 961 GACCACTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
 DB 961 GACCACTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
 QY 1021 CAGATGCTCTCTACCTTCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
 DB 1021 CAGATGCTCTCTACCTTCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
 QY 1081 AAGTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
 DB 1081 AAGTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140  
 QY 1141 AGCGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
 DB 1141 AGCGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
 QY 1201 GATGAAGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
 DB 1201 GATGAAGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
 QY 1261 GAGCGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
 DB 1261 GAGCGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
 QY 1321 ACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380  
 DB 1321 ACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380  
 QY 1381 GCTGCGCTTCCCGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1431  
 DB 1381 GCTGCGCTTCCCGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1431

RESULT 2  
 ABNS2379  
 ID ABNS2379 standard; DNA; 65 BP.  
 AC ABNS2379;  
 DT 15-JUL-2002 (first entry)  
 XX Mouse spliced transcript detection oligonucleotide seq ID NO:25127.  
 DE Human, mouse; rat; splice transcript; detection; RNA transcript;  
 XX splice variant; transcriptome; oligonucleotide library; ss.  
 OS Mus musculus.  
 XX W0200210449-A2.  
 XX PD 07-FEB-2002.

20-JUL-2001; 2001WO-IB01903. *ok AZP alt*  
 XX 28-JUL-2000; 2000US-221607P. *3 too late*  
 PR 02-MAY-2001; 2001US-287724P.  
 XX (COMP-) COMPUGEN INC.  
 PA Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
 PI WPI; 2002-257383/30.  
 DR  
 XX  
 PT New oligonucleotide libraries comprising oligonucleotides which  
 PT selectively hybridize to mRNAs transcribed from a transcription unit of  
 PT a genome, useful for detecting tissue-, pathology-, and  
 PT developmental-specific genes -  
 XX  
 PS Example 1; SEQ ID 25127; 47bp; English.

The present invention describes oligonucleotide libraries for detecting  
 CC messenger RNAs that populate a (sub-)transcriptome, where the  
 CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
 CC transcription units that populate a genome. The library comprises  
 CC several oligonucleotides, each capable of hybridizing selectively to a  
 CC set of messenger RNAs transcribed from a given transcription unit of  
 CC the genome, which encodes one or more messenger RNA splice variants.  
 CC The oligonucleotide libraries are useful for detecting mRNAs from a  
 CC biological sample. In expression profiling studies, in qualitatively or  
 CC quantitatively characterizing the corresponding transcriptome, and in  
 CC detecting RNA transcripts and splice variants of human or animal  
 CC transcriptomes. The libraries may also be used as specialised mini  
 CC libraries to detect transcripts of a sub-transcriptome under a  
 CC particular biological or pathological state, and so allowing the  
 CC detection of tissue- and pathology-specific genes such as those genes  
 CC only expressed in specific tissue under a specific pathological  
 CC condition; to detect developmental specific genes; and to detect RNA  
 CC transcripts and splice variants of a transcriptome of a patient suffering  
 CC from a particular disorder. ABN27253 to ABN59589 represent  
 CC oligonucleotide sequences from rats, humans and mice, which are used in  
 CC the exemplification of the present invention.  
 CC N.B. The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WPI  
 CC at ftp.wip.int/pub/published\_pcl\_sequences.

Sequence 65 BP; 17 A; 13 C; 21 G; 14 T; 0 other;

Query Match 4.5%; Score 65; DB 24; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-21;  
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1114 TATGAGAGCATGAGCAATCTCTGTGAGCGTGGAGCAGCGGAATGTTGGAAC 1173  
 |||||||  
 DB 1 TATGAGAGCATGAGCAATCTCTGTGAGCGTGGAGCAGCGGAATGTTGGAAC 60

OY 1174 ATCGA 1178  
 |||||  
 DB 61 ATCGA 65

RESULT 3  
 AA72433  
 ID AAA72433 standard; cDNA; 1349 BP.

XX AAA72433;  
 AC  
 XX  
 DT 19-DEC-2000 (first entry)

XX Human nucleic acid-binding protein NuABP-52 cDNA.

XX Human nucleic acid-binding protein; NuABP; agonist; antagonist; EST;  
 KM expressed sequence tag; drug screening; recombinant expression; antibody;  
 KM reproductive disorder; infertility; immunological disorder;  
 KM neurological disorder; cell proliferative disorder; cancer; tumour; ss.  
 XX

OS Homo sapiens.

XX WO200044900-A2.

XX 03-AUG-2000. *too late*

PD

PE 28-JAN-2000; 2000WO-US02237. *not applicable under ATPA*

XX 29-JAN-1999; 99US-0117904.

PR 29-JAN-1999; 99US-0117905.

XX (INCY-) INCYTE PHARM INC.

PA Tang YT, Lal P, Hillman JL, Yue H, Azimzai Y, Lu AMD, Baughn MR;  
 PI Tran B, Shih LL, Au-Young JL;

XX WPI; 2000-499332/44.

DR P-PSDB; AAB21048.

XX Novel nucleic acid binding proteins, used to identify agonists and  
 PT antagonists of them, for the treatment of reproductive, immunological,  
 PT neurological and cell proliferative disorders including cancer -

Claim 4; Page 178-179; 180pp; English.

Sequences AAA72382-A72436 represent cDNAs encoding novel human nucleic  
 CC acid-binding proteins (NuABPs; AAB20997-B21051). These cDNAs were  
 CC produced by extension from an appropriate EST (expressed sequence  
 CC tag) using primers designed using the EST. The invention also relates  
 CC to expression constructs, host cells and transgenic organisms comprising  
 CC a human NuABP nucleic acid, recombinant production of the human NuABPs,  
 CC and antibodies against the human NuABPs, and also to methods of  
 CC screening modulators of human NuABP activity or expression. The human  
 CC NuABPs, and their agonists and antagonists are used to treat diseases  
 CC associated with overexpression or underexpression of functional NuABPs.  
 CC Human NuABP proteins and nucleotides, and NuABP agonists and antagonists  
 CC can be used to diagnose, treat and prevent reproductive, immunological,  
 CC neurological and cell proliferative disorders. Reproductive disorders  
 CC that may be treated using compositions of the invention include  
 CC infertility, endometriosis, disruptions of the menstrual cycle and  
 CC disruptions of spermatogenesis. Immunological disorders that may be  
 CC treated include AIDS, allergies, and autoimmune disorders such as  
 CC multiple sclerosis, rheumatoid arthritis, diabetes and systemic lupus  
 CC erythematosus. Neurological disorders that may be treated include  
 CC epilepsy, neurodegenerative conditions such as Alzheimer's disease and  
 CC Parkinson's disease, prion diseases such as Creutzfeldt-Jakob disease,  
 CC and mental disorders such as schizophrenia. Cell proliferative disorders  
 CC that may be treated include a wide variety of cancers, and also  
 CC arteriosclerosis, atherosclerosis, cirrhosis and psoriasis.

Sequence 1349 BP; 328 A; 372 C; 428 G; 221 T; 0 other;

Query Match 2.5%; Score 36; DB 21; Length 1349;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;  
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1099 CGGCGAGCCAGCGTATGAGACATGAGCAATTC 1134  
 |||||||  
 DB 1029 CGGCGAGCCAGCGTATGAGACATGAGCAATTC 1064

RESULT 4  
 AAH68563  
 ID AAH68563 standard; cDNA; 1913 BP.

XX AAH68563;  
 AC  
 XX  
 DT 13-SEP-2001 (first entry)

XX Human protein HP03115 coding sequence.

XX Human; gene therapy; tumour; ss.  
 XX

OS Homo sapiens.  
PN WO200142302-A1.  
XX  
XX  
PD 14-JUN-2001. ~~XXXXXXXXXX~~ *OK Absolute but*  
XX *pubshw@*  
PF 06-DEC-2000; 2000WO-JP08631. *not in English*  
PR 06-DEC-1999; 99JP-0346863  
PR 06-DEC-1999; 99JP-0346864  
PR 08-FEB-2000; 2000JP-0031062  
PR 10-FEB-2000; 2000JP-0034090  
PR 10-FEB-2000; 2000JP-0034091  
PR 14-FEB-2000; 2000JP-0035829  
PR 14-FEB-2000; 2000JP-0035899  
PR 14-MAR-2000; 2000JP-0071161  
PR 30-MAY-2000; 2000JP-0160851.  
XX  
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
XX  
XX Kato S, Eguchi C, Saeki M:  
XX  
XX WPI: 2001-381646/40.  
XX P-PSDB; AMG93278.  
XX  
XX Human protein originated from tumor cell line, applicable as drug,  
XX reagent for studying intracellular protein networks and protein source  
XX for drug screening, also encoded cDNA for gene diagnosis and gene  
XX therapy -  
XX  
XX Claim 3; Pages 236-240; 471pp; Japanese.  
XX  
XX  
XX The present sequence is a human protein coding sequence. The human  
XX protein, preferably originated from tumor cell line, is applicable as a  
XX drug, a reagent for studying intracellular protein networks and a protein  
XX source for screening proteins for binding low molecular weight drugs. The  
XX human protein coding sequence is useful for gene diagnosis and gene  
XX therapy, expression vectors and transformant cells for detection of  
XX ligands and receptors.  
XX  
XX Sequence 1913 BP; 446 A; 537 C; 612 G; 318 T; 0 other;  
XX  
XX  
XX Query Match 2.5%; Score 36; DB 22; Length 1913;  
XX Best Local Similarity 100.0%; Pred. No. 2.6e-07;  
XX Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX QY 1099 GGGCCGAGCAGGCTATGAGAGCATGGAGCAATTC 1134  
XX ||||||||||||||||||||||||||||||||||||||||  
XX Db 1203 GGGCCGAGCAGGCTATGAGAGCATGGAGCAATTC 1238  
XX  
XX  
XX RESULT 5  
XX AAH78026  
XX ID AAH78026 standard; cDNA; 1990 BP.  
XX  
XX AAH78026;  
XX  
XX 26-NOV-2001 (first entry)  
XX  
XX Nucleotide sequence of human cardiomyopathy associated protein (CAP).  
XX  
XX  
XX Human; cardiomyopathy associated protein; CAP; myocardial biopsy;  
XX left ventricle assist device; LVAD; cardiomyopathy;  
XX ventricular tachyrythmia; ss.  
XX  
XX Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
XX FH CDS /tag= a  
XX FT /product= "cardiomyopathy associated protein (CAP)"  
XX  
XX  
XX MO200162767-A1.

```

XX XX    30-AUG-2001.
PD PF    26-FEB-2001; 2001MO-USO5888. OK Ser HZPA
XX XX    24-FEB-2000; 2000US-0184825. OK Ser AIDA
PR PR    (GENE-) GENE LOGIC INC.
XX PA
XX PA    Bedarik D, Greene J, White M;
DR DR    MPI; 2001-570621/64.
XX XX    P-PADB; AAG63832.
XX XX
PT PT    Nucleic acid encoding a cardiomyopathy associated protein that is
PT PT    differentially expressed in human left ventricle assist device (LVAD)
PT PT    myocardial biopsy samples, useful for diagnosing, preventing or
PT PT    treating cardiomyopathy -
PS PS    Claim 2; Page 49-51; 55pp; English.
XX XX
CC CC    The present sequence encodes a human cardiomyopathy associated
CC CC    protein (CAP) that is differentially expressed in human left ventricle
CC CC    assist device (LVAD) myocardial biopsy samples. The predicted
CC CC    isolectric point of CAP is approximately 12.2. The CAP polynucleotide
CC CC    and polypeptide are useful for diagnosing, preventing or treating
CC CC    cardiomyopathy. The treatment diminishes the occurrence of at least
CC CC    one of the following symptoms associated with cardiomyopathy: reduced
CC CC    ejection fraction, increased left ventricular diastolic dimension,
CC CC    decreased ventricular wall thickness, increased atrial size, valvular
CC CC    regurgitation, exertional intolerance or ventricular tachyrrhythmia.
CC CC    Modulators of CAP are also useful for preventing or treating
XX XX    cardiomyopathy.
SQ SQ    Sequence 1990 BP; 524 A; 535 C; 613 G; 318 T; 0 other;

Query Match      2.5%; Score 36; DB 22; Length 1990;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches   36; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

QY 1099 CGCGGAGGCCAGGGCTATGAGAGCATGAGCAATTTC 1134
    |||||
DB 1183 CGGCGGAGGCCAGGCTATGAGAGCATGAGCATTTC 1218

RESULT 6
ABN85313
ID ID ABN85313 standard; cDNA; 2040 BP.
AC AC
XX XX ABN85313;
DT DT
XX XX 30-SEP-2002 (first entry)
DE DE Human cytoskeleton-associated protein, CSAP-4, coding sequence.
XX XX
XX XX Human; cytoskeleton-associated protein; CSAP; CSAP-4;
KW KW cell proliferative disorder; viral infection; neurological disorder;
KM KM transgenic animal; antihypertensive; antisporadic; antiinflammatory;
KV KV vinclozole; anticonvulsant; vasodilator; cerebroprotective; neurotropic;
XX XX neuroprotective; cytoskeletal; gene therapy; gene; ss.
OS OS Homo sapiens.
XX XX
XX XX Key Location/Qualifiers
FT FT CDS 291..1271
FT FT /*tag= a
FT FT /product= "CSAP-4"
PN PN WO200253719-A2.
PD PD 11-JUL-2002.
FF FF 04-JAN-2002; 2002WO-US001178. OK BT OA / F

```

XX 04-JAN-2001; 2001US-260085P. } *not too late*  
 PR 13-FEB-2001; 2001US-268554P.  
 PR 14-FEB-2001; 2001US-269111P.  
 PR 23-FEB-2001; 2001US-271211P.  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA Lu DAM, Baughn MR, Yao MG, Ding L, Honchell CD, Yue H, Tang YT;  
 PI Warren BA, Duggan BM, Xu Y, Walla NK, Griffin JA, Stewart EA;  
 PI Gandhi AR, Khan FA, Thangavelu K, Ison CH, Azimzai Y, Hatfalia AJA;  
 PI Gietzen KJ, Lal PG, Sanjanwala MM, Elliott VS;  
 XX WPI: 2002-583611/62.  
 DR P-PSDB; ABB83475.  
 XX Novel isolated human cytoskeleton-associated protein for diagnosing,  
 PT treating or preventing atherosclerosis, psoriasis, leukemia, epilepsy,  
 PT ischemic cerebrovascular disease, cerebral neoplasms and Alzheimer's  
 PT disease -  
 XX Claim 5; Page 154-155; 167pp; English.  
 CC The present sequence is the coding sequence for a human  
 CC cytoskeleton-associated protein (CSAP). The CSAP and its coding sequence  
 CC are useful in the diagnosis, treatment and prevention of a cell  
 CC proliferative disorder such as actinic keratosis, atherosclerosis,  
 CC psoriasis, primary thrombocythaemia, leukaemia; a viral infection such as  
 CC those caused by adenoviruses (acute respiratory disease, pneumonia),  
 CC arenaviruses (lymphocytic choriomeningitis); and a neurological disorder  
 CC such as epilepsy, ischaemic cerebrovascular disease, stroke, cerebral  
 CC neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease or  
 CC amyotrophic lateral sclerosis. The CSAP coding sequence is also useful  
 CC for creating knock out or knock in humanised animals or transgenic  
 CC animals to model human diseases.  
 XX Sequence 2040 BP; 465 A; 584 C; 650 G; 341 T; 0 other;  
 SO  
 XX Query Match 2.5%; Score 36; DB 24; Length 2040;  
 XX Best Local Similarity 100.0%; Pred. No. 2.6e-07;  
 XX Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1099 CGCGGAGCCAGCTATGAGCATGAGCAATTC 1134  
 DB 1095 CGCGGAGCCAGCTATGAGCATGAGCAATTC 1130  
 RESULT 7  
 ID AAF27653 standard; DNA; 1500 BP.  
 AC AAF27653;  
 XX 30-MAR-2001 (first entry)  
 DE DNA encoding human transcriptional regulator protein #22.  
 XX Human; transcriptional regulator protein; TXREG; ds.  
 XX Homo sapiens.  
 XX MO200078954-A2.  
 XX PD 28-DEC-2000. *too late*  
 XX PF 15-JUN-2000; 2000MO-US16766. *Not available under ATPA*  
 XX PR 18-JUN-1999; 99US-0140109.  
 XX PA (INCY-) INCYTE GENOMICS INC.  
 XX PI Lal P, Yue H, Tang YT, Baughn MR, Azimzai Y, Tran B;

DR WPI: 2001-041425/05.  
 XX Isolated polypeptide with a human transcriptional regulator protein  
 PT sequence is useful for the diagnosis, prevention and treatment of  
 PT disorders associated with the immune, reproductive and cardiovascular  
 PT systems -  
 XX Claim 5; Page 135-136; 142pp; English.  
 CC The present invention relates to human transcriptional regulator  
 CC protein (TXREG) sequences. The antagonist and an agonist of the proteins  
 CC of the invention are used to treat disorders associated with decreased  
 CC or increased expression or activity of TXREG.  
 XX Sequence 1500 BP; 401 A; 378 C; 439 G; 282 T; 0 other;  
 SO  
 XX Query Match 2.4%; Score 35; DB 22; Length 1500;  
 XX Best Local Similarity 100.0%; Pred. No. 8.1e-07;  
 XX Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 685 CTGCCACCATTTACAAACGCCAGAGAGTGAAGT 719  
 DB 777 CTGCCACCATTTACAAACGCCAGAGAGTGAAGT 811  
 RESULT 8  
 ID AAC75760 standard; CDNA; 650 BP.  
 AC AAC75760;  
 XX 08-FEB-2001 (first entry)  
 DE Human ORFX ORF1315 polynucleotide sequence SEQ ID NO:2629.  
 XX  
 XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnery; antiproliferative; antiparkinsonian; neurotropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive; ss.  
 XX Homo sapiens.  
 XX MO200058473-A2.  
 XX PD 05-OCT-2000. *too late*  
 XX PF 31-MAR-2000; 2000MO-US08621. *Not available under ATPA*  
 XX PR 31-MAR-1999; 99US-0127607.  
 XX PR 02-APR-1999; 99US-0127636.  
 XX PR 05-APR-1999; 99US-0127728.  
 XX PR 30-MAR-2000; 2000US-0540763.  
 XX PA (CURA-) CURAGEN CORP.  
 XX PI Shimkets RA, Leach M;  
 XX WPI: 2000-602362/57.  
 XX P-PSDB; AAB41551.  
 XX Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -

XX Claim 5; Page 1873; 5507pp; English.

PS AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC sequences represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;  
 CC antiproliferic; antiparkinsonian; nootropic; neuroprotective;  
 CC osteoprotic; anticonvulsant; antiarthritic; immunosuppressive;  
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotrophic;  
 CC antidiabetic; hypotensive; dermatologic; immunosuppressive;  
 CC antitumor; antibacterial; antiviral; antifungal; antineumatic;  
 CC antihypertoid; and antianemic. The sequences can be used for determining  
 CC the presence of or predisposition to, or preventing or treating  
 CC pathological conditions associated with an ORFX-associated disorder. The  
 CC nucleic acids can be used to express ORFX proteins in gene therapy  
 CC vectors. The proteins and nucleic acids may be used to treat cancers,  
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

SQ Sequence 650 BP; 171 A; 150 C; 212 G; 116 T; 1 other;

Query Match 2.0%; Score 29; DB 21; Length 650;  
 Best Local Similarity 100.0%; Pred. No. 0.00072;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 GTGGTTTCAAGCCGCTCTAGGCGATGC 239  
 DB 85 GTGGTTTCAAGCCGCTCTAGGCGATGC 113

RESULT 9  
 ID ABA99063 standard; DNA; 1597 BP.

AC ABA99063;  
 XX 15-JUL-2002 (first entry)  
 DE Murine muscle ring finger protein 3 (MURF-3) coding sequence.  
 XX  
 DE Muscle ring finger; MURF-3; mouse; cardiant; microtubule;  
 KW intermediate filament; striated muscle; cardiac hypertrophy;  
 KW heart disease; gene; ds.  
 XX  
 OS Mus musculus.

XX Key Location/Qualifiers  
 FT 299..1330  
 FT /tag= a  
 FT /product= "MURF-3"

XX WO200206318-A2.

XX 24-JAN-2002.

XX 18-JUL-2001; 2001WO-US22896.

XX 18-JUL-2000; 2000US-219020P.

XX (TEXA ) UNIV TEXAS SYSTEM

XX Olson EN, Spencer JA;

XX WPI; 2002-241506/29.

XX P-PSDB; ABB08277.

XX Novel muscle ring finger protein useful for drug screening, and for

PT diagnosing and treating diseases, particularly cardiomyopathies  
 XX Claim 4; Page 131-133; 134pp; English.

XX The sequence encodes murine muscle ring finger protein 3 (MURF-3). The  
 XX invention relates to a purified muscle ring finger (MURF) protein,  
 CC selected from MURF-1, MURF-2 and MURF-3. The MURF proteins of the  
 CC invention are involved in microtubule and intermediate filament  
 CC stabilisation of striated muscle cells and have cardiant activity.  
 CC The MURF proteins are useful for screening a candidate substance  
 CC for MURF protein-binding activity. In a cell, cell-free system or  
 CC in vivo, and its effect on interaction of MURF with microtubules,  
 CC homodimerisation of MURF, MURF-1, MURF-2 or MURF-3 stabilisation  
 CC of microtubules, interaction of MURF with intermediate filament  
 CC e.g. desmin, vimentin and cytokeratin, and heterodimerisation of MURF.  
 CC The screened compounds are useful for treating and preventing  
 CC cardiac hypertrophy and heart diseases. MURF proteins are useful as  
 CC antigens to immunise animals for the production of antibodies.

SQ Sequence 1597 BP; 405 A; 404 C; 438 G; 350 T; 0 other;

Query Match 2.0%; Score 28; DB 24; Length 1597;  
 Best Local Similarity 100.0%; Pred. No. 0.0022;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 505 AACATCATTTGACATCTACACAGAGAGT 532  
 DB 596 AACATCATTTGACATCTACACAGAGAGT 623

RESULT 10  
 ID AAS26314 standard; cDNA; 587 BP.

AC AAS26314;  
 XX 07-NOV-2001 (first entry)

DE Human cDNA encoding a novel secreted protein, Seq ID 493.

XX Human; immunosuppressive; antiarthritic; ss; antirheumatic;  
 KW cytosolic; cardiant; vasotropic; cerebroprotective; nootropic;  
 KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
 KW vulnery; secreted protein; rheumatoid arthritis;  
 KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;  
 KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;  
 KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;  
 KW corneal infection; wound healing; epithelial cell proliferation;  
 KW skin ageing; food additive; preservative; antiproliferative.

XX Homo sapiens.

XX WO200155322-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01341.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

*ORFX A data  
 but Belkumit is better*

PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226686.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251088.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 PR  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI WPI: 2001-488783/53.  
 XX P-PSDB; AAT016327.  
 DR  
 XX  
 PT New nucleic acid molecules encoding 461 human secreted proteins for  
 PT diagnosing, preventing, treating or ameliorating medical conditions and  
 PT used as food additives or preservatives -  
 PT  
 PT  
 XX  
 PS  
 PS Claim 1; SEQ ID No 493; 980pp; English.  
 XX  
 XX The invention relates to isolated nucleic acid molecules and their  
 CC encoded secreted proteins. The nucleic acids and proteins are used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They  
 CC are also used in diagnosing a pathological condition or susceptibility  
 CC to a pathological condition. Antibodies to the proteins can also  
 CC be used in alleviating symptoms associated with the disorders and in  
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
 CC immunoassay assays (ELISA). Disorders which are diagnosed or treated  
 CC include autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,

CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi  
CC and ocular disorders e.g. corneal infection, and many other  
CC disorders listed in the specification. The polypeptides can also  
CC be used to aid wound healing and epithelial cell proliferation, to  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
CC minerals, cofactors and other nutritional components. The present  
CC sequence encodes a novel secreted protein of the invention.

Query Match 1.9%; Score 27; DB 22; Length 587;

Best Local Similarity 100.0%; Pred. No. 0.0069;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 508 ATCATGTGACATCTACAGCAGAGTCC 534  
Db 445 ATCATGTGACATCTACAGCAGAGTCC 471

## RESULT 11

AAS25855 ID AAS25855 standard; cDNA: 1762 BP.

XX AAS25855;

DT 07-NOV-2001 (first entry)

DE Human cDNA encoding a novel secreted protein, Seq ID 34.

XX Human; immunosuppressive; antiarthritic; ss; antihemetic;  
XX cytoskeletal; cardiac; vasotropic; cerebroprotective; nocotropic;  
XX neuroprotective; antibacterial; virucide; fungicide; optalmallogical;  
XX vulnary; secreted protein; rheumatoid arthritis;  
XX hyperproliferative disorder; cardiovascular disorder; cardiac arrest;  
XX cerebrovascular disorder; cerebral ischaemia; angiogenesis;  
XX nervous system disorder; Alzheimer's disease; infection; ocular disorder;  
XX corneal infection; wound healing; epithelial cell proliferation;  
XX skin ageing; food additive; preservative; antiproliferative.

OS Homo sapiens.

PN WO20015322-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01341.

XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 12-SEP-2000; 2000US-0232081.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 20-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241806.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246479.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.





Db 316 GTGGCTTCAAGCCGCTCTAGCGGA 341

## RESULT 13

ABL24238/C  
ID ABL24238 standard; DNA; 2620 BP.

XX ABL24238;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 24187.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ds.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Claim 1; SEQ ID NO 24187; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

XX sequences (AB101840-AB16175) and the encoded proteins

XX (ABBS7737-ABW2072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published\_pcl\_sequences.

XX Sequence 2620 BP; 625 A; 598 C; 595 G; 802 T; 0 other;

XX Query Match 1.4%; Score 20; DB 23; Length 2620;

XX Best Local Similarity 100.0%; Pred. No. 18;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 472 GTCTATGCGCTGACGCGAA 491

XX 985 GTCTATGCGCTGACGCGAA 966

XX Db

XX RESULT 14

XX ABL14743

XX ID ABL14743 standard; CDNA; 8574 BP.

XX ABL14743;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 38711.

KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-PADB; ABB70640.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Claim 1; SEQ ID NO 38711; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

XX sequences (AB101840-AB16175) and the encoded proteins

XX (ABBS7737-ABW2072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published\_pcl\_sequences.

XX Sequence 8574 BP; 1935 A; 2333 C; 2424 G; 1882 T; 0 other;

XX Query Match 1.4%; Score 20; DB 23; Length 8574;

XX Best Local Similarity 100.0%; Pred. No. 17;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 872 AGGAGCGCAAGCGGCAACTG 891

XX 3041 AGGAGCGCAAGCGGCAACTG 3060

XX Db

XX RESULT 15

XX ABL14742/C

XX ID ABL14742 standard; CDNA; 13239 BP.

XX ABL14742;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 38708.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE ) PE CORP NY.  
 PT Venter JC, Adams M, Li PWD, Myers EW;  
 PI  
 XX WPI: 2001-6556860/75.  
 DR P-PDB; ABB70639.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 CC genes from Drosophila and for elucidating cell signalling and cell-cell  
 CC interactions -  
 XX  
 CC Claim 1; SEQ ID NO 38708; 21bp + Sequence Listing; English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins  
 CC (AB857737-AB872072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX Sequence 13239 BP; 3468 A; 3272 C; 3140 G; 3359 T; 0 other;  
 SQ  
 Query Match 1.4%; Score 20; DB 23; Length 13239;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 872 AGGACGCAAGGCGCACTG 891  
 DB 7067 AGGACGCAAGGCGCACTG 7048  
 ||||||||||||||||||||  
 RESULT 16  
 ID ABR62174/c  
 AC ABR62174 standard; cDNA; 488 BP.  
 XX  
 AC ABR62174;  
 DT 18-JUN-2002 (first entry)  
 XX  
 DE Rat sequence differentially expressed in response to a hepatotoxin #81.  
 XX  
 KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;  
 KW differential expression; centrilobular necrosis; steatosis.  
 XX  
 OS Rattus norvegicus.  
 XX  
 PN WO200210453-A2.  
 PD 07-FEB-2002.  
 XX  
 PE 30-JUL-2001; 2001WO-US23872.  
 XX  
 PR 31-JUL-2000; 2000US-222040P.  
 PR 02-NOV-2000; 2000US-244880P.  
 PR 11-MAY-2001; 2001US-290029P.  
 PR 15-MAY-2001; 2001US-290645P.  
 PR 22-MAY-2001; 2001US-292336P.  
 PR 06-JUN-2001; 2001US-295798P.  
 PR 13-JUN-2001; 2001US-297457P.  
 PR 19-JUN-2001; 2001US-298884P.  
 PR 09-JUL-2001; 2001US-303459P.  
 XX  
 PA (GENE-) GENE LOGIC INC.  
 XX  
 PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;  
 XX WPI: 2002-241625/29.

XX  
 PT predicting toxic effects of compounds or the progression of these toxic  
 PT effects by determining the changes in gene expression in tissues or  
 PT cells exposed to the toxin and comparing these to gene expression in  
 PT unexposed tissues or cells -  
 XX  
 XX Claim 1; Seq ID No 81; 239bp; English.  
 XX  
 CC The invention relates to methods for predicting toxic effects of  
 CC compounds or the progression of these toxic effects by determining the  
 CC global changes in gene expression in tissues or cells exposed to the  
 CC toxin and comparing these to gene expression in unexposed tissues or  
 CC cells. Also included are methods of predicting at least one toxic  
 CC effect of a compound or progression of a toxic effect, preferably the  
 CC hepatotoxicity of a compound, comprising detecting the level of  
 CC expression in a tissue or cell sample exposed to the compound of two or  
 CC more genes listed in the specification, where differential expression of  
 CC the genes is indicative of at least one toxic effect or progression.  
 CC The method can also be used to identify an agent which modulates the  
 CC toxic response and predict cellular pathways that a compound modulates  
 CC in a cell. The methods utilise a set of at least two probes (on a solid  
 CC support in kit form), where each of the probes comprises a sequence that  
 CC specifically hybridises to a gene listed in the specification, a computer  
 CC system comprising a database containing information identifying the  
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a  
 CC set of genes comprising at least two genes listed in the specification,  
 CC and a user interface to view the information used to present information,  
 CC identifying the expression level in a tissue or cell of at least one gene  
 CC listed in the specification. The method is useful for elucidating global  
 CC changes in gene expression and for identifying toxicity markers in  
 CC tissues or cell exposed to a known toxin. The genes may be used as  
 CC toxicity markers in drug screening and toxicity assays. The genes and  
 CC gene expression information may be used as diagnostic markers for the  
 CC prediction or identification of the physiological state of tissue or cell  
 CC sample that has been exposed to a compound or agent. Hepatotoxicity  
 CC is characterised by centrilobular necrosis and steatosis. The present  
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene  
 CC which is differentially expressed in response to a hepatotoxic agent.  
 CC  
 XX  
 SQ Sequence 488 BP; 93 A; 136 C; 136 G; 123 T; 0 other;  
 Query Match 1.3%; Score 19; DB 24; Length 488;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1197 TGGGATGAGAGATGAC 1215  
 DB 390 TGGGATGAGAGATGAC 372  
 ||||||||||||||||||||  
 RESULT 17  
 ID ABR34528  
 AC ABR34528 standard; cDNA; 596 BP.  
 XX  
 AC ABR34528;  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE Human cDNA for novel secreted protein, SPQ ID 297.  
 XX  
 KW Human; ss; gene; secreted protein; immune deficiency; viral infection;  
 KW bacterial infection; fungal infection; autoimmune disorder; burn;  
 KW rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;  
 KW diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;  
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; tumour;  
 KW coagulation disorder; hemophilia; inflammatory disorder; ulcer;  
 KW tissue regeneration; wound healing; haematopoiesis; myeloid deficiency;  
 KW lymphoid cell deficiency.  
 XX  
 OS Homo sapiens.  
 XX  
 PI WO200177290-A2.  
 XX

PD 18-OCT-2001.  
 XX  
 PF 29-MAR-2001; 2001WO-US10295.  
 XX  
 PR 06-APR-2000; 2000US-194941P.  
 XX  
 PA (GENY ) GENETICS INST INC.  
 XX  
 PI Wong GG, Clark HF, Feschel K, Agostino MJ, Howes SH, Resnick RJ,  
 PI Gulukota K, Graham JR;  
 XX  
 DR WPI: 2002-179323/23.  
 XX  
 PT Six hundred and twenty five polynucleotides derived from a variety of  
 PT human tissue sources which encode secreted proteins, useful for  
 PT treating immune deficiencies and disorders such as autoimmune disorders  
 PT  
 PS Claim 1; Page 152; 339pp; English.  
 XX  
 CC The invention relates to 625 polynucleotides which have been derived from  
 CC a variety of human tissue sources and which encode novel secreted  
 CC proteins, their complements and sequences that hybridise to them.  
 CC Also included are a vector comprising the polynucleotide, a host cell  
 CC transformed with the vector, the proteins encoded by the  
 CC polynucleotides, antibodies that bind to the proteins and identification  
 CC of modulators of the proteins or the expression of the polynucleotide.  
 CC The polynucleotides can be used as probes for the identification  
 CC and isolation of full length cDNA and genomic DNA. The polynucleotides  
 CC and proteins can also be used as nutritional supplements. The protein  
 CC is useful in the treatment of various immune deficiencies and disorders  
 CC such as viral infections, bacterial infections, fungal infections,  
 CC autoimmune disorders (e.g. rheumatoid arthritis, multiple sclerosis,  
 CC conditions (e.g. asthma). They are also useful for treating  
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's  
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),  
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also  
 CC useful for tissue regeneration, for wound healing and in the treatment  
 CC of burns, incisions and ulcers. The proteins are also useful for  
 CC regulating haematopoiesis, for treating myeloid or lymphoid cell  
 CC deficiencies. The present sequence is one of the 625 cDNA sequences  
 CC encoding a secreted protein.  
 XX  
 SQ Sequence 596 BP; 116 A; 152 C; 235 G; 93 T; 0 other;  
 XX  
 Query Match 1.3%; Score 19; DB 24; Length 596;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1247 CGGGCGTGGAGGAGGACG 1265  
 ||||||||||||||||  
 DB 250 CGGGCGTGGAGGAGGACG 268  
 RESULT 18  
 ABA82018/C  
 ID ABA82018 standard; cDNA; 629 BP.  
 XX  
 AC ABA82018;  
 XX  
 DT 28-JUN-2002 (first entry)  
 XX  
 DE Human cancer related polynucleotide SEQ ID NO 408.  
 XX  
 KW Human; cytosolic; gene expression; gene mapping; tissue profiling;  
 KW gene therapy; cancer; tumour; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PA WO200214500-A2.  
 XX  
 PI 21-FEB-2002.  
 XX  
 DR

XX  
 PF 16-AUG-2001; 2001WO-US25840.  
 XX  
 PR 16-AUG-2000; 2000US-226326P.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 PA (HXSE-) HXSED INC.  
 XX  
 PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F,  
 PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;  
 XX  
 DR WPI: 2002-241905/29.  
 XX  
 PT New nucleic acid for producing a polypeptide, detecting differentially  
 PT expressed genes correlated with a cancerous state of a mammalian cell,  
 PT and inhibiting tumor growth -  
 PT  
 PS Claim 1; SEQ ID NO 408; 883pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated polynucleotide (ABA827253-ABA833262)  
 CC with cytosolic activity. The polynucleotide is used to produce a  
 CC polypeptide, to detect differentially expressed genes correlated with a  
 CC cancerous state of a mammalian cell and to inhibit tumour growth. The  
 CC polynucleotide is used as a probe in mapping and tissue profiling. The  
 CC encoded polypeptide and antibodies to the polypeptide can also be used  
 CC for therapeutic and diagnostic purposes. The polynucleotide is useful for  
 CC gene therapy.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 629 BP; 172 A; 195 C; 90 G; 172 T; 0 other;  
 XX  
 Query Match 1.3%; Score 19; DB 24; Length 629;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1080 GAAGGTGAGCTGCGACGA 1098  
 ||||||||||||||||  
 DB 253 GAAGGTGAGCTGCGACGA 235  
 RESULT 19  
 ABA82018/C  
 ID ABA82018 standard; cDNA; 637 BP.  
 XX  
 AC ABA82018;  
 XX  
 DT 25-JAN-2002 (first entry)  
 XX  
 DE Human wound healing related polynucleotide SEQ ID NO 97.  
 XX  
 KW Human; mouse; vulnery; dermatological; skin disorder; wound healing;  
 KW gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PA CA2325226-A1.  
 XX  
 PD 17-MAY-2001.  
 XX  
 PF 16-NOV-2000; 2000CA-2325226.  
 XX  
 PR 17-NOV-1999; 99DE-1055349.  
 PR 17-DEC-1999; 99US-0172511.  
 PR 20-JUN-2000; 2000DE-1030149.  
 XX  
 PA (SWIT-) SWITCH BIOTECH AG.  
 XX  
 PI Regenbogen J, Wolf E, Goppelt A, Werner S, Halle J;  
 XX  
 DR WPI: 2001-433142/47.  
 XX

PT Use of novel polypeptide or its variant or nucleic acid encoding the  
 PT polypeptide for diagnosing and/or preventing and/or treating skin  
 PT disorders and/or treatment in wound healing or for identifying active  
 PT substances -  
 PS Disclosure; Page 243-244; 265pp; English.  
 XX  
 CC The invention relates to the use of a polypeptide (ABB44544-ABB44601,  
 CC ABB44606-ABB44623) or its variant or encoding nucleic acid  
 CC (ABA81990-ABA81995, ABA82016-ABA82032) with vulnerary and/or  
 CC dermatological activity for the diagnosis, prevention and treatment of  
 CC skin disorders and treatment in wound healing or for the identification  
 CC of pharmacologically active substances. The nucleic acids are useful in  
 CC gene therapy.  
 CC Note: The printed sequence listing for this specification was incomplete,  
 CC terminating part way through SEQ ID NO 106. The remaining data was  
 CC obtained from EPO data for an equivalent patent (EP1114862).  
 XX  
 SQ Sequence 637 BP; 131 A; 184 C; 127 G; 195 T; 0 other;  
 Query Match 1.3%; Score 19; DB 22; Length 637;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 908 GGGAGCAGGAGGAGCAATT 926  
 Db 62 GGGAGCAGGAGGAGCAATT 44  
 RESULT 20  
 AAS22581  
 ID AAS22581 standard; cDNA; 737 BP.  
 XX  
 AC AAS22581;  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human cDNA encoding a novel human protein #147.  
 XX  
 KW Human; novel protein; ss; Antinaemic; osteopathic; antiinflammatory;  
 KW immunomodulatory; cytosolic; neuroprotective; vulnerary; nootropic;  
 KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;  
 KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;  
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
 KW tissue regeneration; immune disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155437-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 25-JAN-2001; 2001WO-US02623.  
 XX  
 PR 25-JAN-2000; 2000US-0491404.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Dirmannac RT;  
 XX  
 DR WPI: 2001-451939/48.  
 DR P-PSDB: AAU14276.  
 XX  
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,  
 PT nervous system disorders, and for regenerating bone and cartilage -  
 PS Claim 1; Page 375-376; 894pp; English.  
 XX  
 CC The invention relates to polynucleotides encoding novel human  
 CC proteins or their active domains. The polypeptides, polynucleotides and  
 CC antibodies raised against the polypeptides are used in a method of  
 CC treatment of a mammal and prevention of disorders caused by the aberrant

CC protein expression or activity. The polypeptides can be used as  
 CC molecular weight markers, food supplements, and in antibody production.  
 CC The polypeptides are used to identify compounds which bind to the  
 CC polypeptides. Polynucleotides of the invention are used as probes and  
 CC primers, for sequencing, for chromosome or gene mapping, in the  
 CC production of recombinant proteins, and in generating anti-sense DNA or  
 CC RNA and in gene therapy. Polypeptides of the invention can be used to  
 CC target drugs to a tumour, in assays to determine biological activity, to  
 CC raise antibodies/elicit an immune response, to determine quantitative  
 CC protein levels, as tissue markers, and to isolate receptors or ligands.  
 CC Polypeptides of the invention may also be useful in treating platelet  
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
 CC the proliferation, differentiation and survival of stem cells, as a  
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,  
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,  
 CC anti-inflammatory diseases, nervous system disorders, and infection.  
 CC The present sequence encodes a protein of the invention.  
 XX  
 SQ Sequence 737 BP; 246 A; 111 C; 146 G; 234 T; 0 other;  
 Query Match 1.3%; Score 19; DB 22; Length 737;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1115 ATGAGAGCATGAGCAATT 1133  
 Db 699 ATGAGAGCATGAGCAATT 717

Search completed: December 3, 2002, 10:41:33  
 Job time : 309 secs

47 C  
D  
2.4  
E  
F  
G  
A  
B  
47

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 10:41:43 ; Search time 65 Seconds  
(without alignments)  
8478.465 Million cell updates/sec

Title: US-09-908-988B-1

Perfect score: 1431  
Sequence: 1 aagagagtgtacagagagtgtc.....ataagagactcaagtgtccc 1431

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 341543 seqs, 192557720 residues

Word size : 15

Total number of hits satisfying chosen parameters: 832

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database :

Published Applications - NA: \*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq: \*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq: \*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq: \*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq: \*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq: \*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1431	100.0	1431	10	US-09-908-988B-1
2	28	2.0	1597	10	US-09-908-988B-5
3	27	1.9	587	10	US-09-764-864-493
4	27	1.9	1762	10	US-09-764-864-34
5	19	1.3	488	10	US-09-917-800A-81
6	19	1.3	2513	10	US-09-822-849A-268
7	18	1.3	399	10	US-09-960-352-207
8	18	1.3	444	10	US-09-960-352-13152
9	18	1.3	445	10	US-09-917-800A-577
10	18	1.3	467	9	US-09-736-457-753
11	18	1.3	467	9	US-09-902-941-753
12	18	1.3	746	10	US-09-822-849A-580
13	18	1.3	860	10	US-09-833-381-303
14	18	1.3	3033	12	US-10-044-090-508
15	18	1.3	4153	10	US-09-917-800A-1593
16	18	1.3	684973	10	US-09-263-959-1
17	17	1.2	133	10	US-09-776-695-22
18	17	1.2	192	10	US-09-878-574-9448
19	17	1.2	230	10	US-09-864-761-22504

20	17	1.2	231	10	US-09-923-876-2132	Sequence 2132, Ap
21	17	1.2	262	10	US-09-878-574-8578	Sequence 8578, Ap
22	17	1.2	264	10	US-09-294-093B-2808	Sequence 2808, Ap
23	17	1.2	264	10	US-09-764-864-499	Sequence 499, Ap
24	17	1.2	274	10	US-09-923-876-5150	Sequence 5150, Ap
25	17	1.2	286	10	US-09-878-574-10927	Sequence 10927, A
26	17	1.2	292	10	US-09-294-093B-1300	Sequence 1300, Ap
27	17	1.2	295	10	US-09-294-093B-1812	Sequence 1812, Ap
28	17	1.2	347	10	US-09-864-761-22849	Sequence 22849, A
29	17	1.2	370	10	US-09-960-352-2716	Sequence 2716, Ap
30	17	1.2	417	10	US-09-867-701-4631	Sequence 4631, Ap
31	17	1.2	446	10	US-09-764-864-482	Sequence 482, Ap
32	17	1.2	459	10	US-09-864-761-31683	Sequence 31683, A
33	17	1.2	465	10	US-09-864-761-6101	Sequence 6101, Ap
34	17	1.2	485	10	US-09-864-761-5742	Sequence 5742, Ap
35	17	1.2	498	10	US-09-864-761-15156	Sequence 15156, A
36	17	1.2	497	10	US-09-783-590-8999	Sequence 8999, Ap
37	17	1.2	535	10	US-09-974-300-3014	Sequence 3014, Ap
38	17	1.2	570	10	US-09-864-761-14784	Sequence 14784, A
39	17	1.2	588	9	US-09-768-020-75	Sequence 75, Appl
40	17	1.2	774	10	US-09-764-864-40	Sequence 40, Appl
41	17	1.2	815	10	US-09-818-954A-18	Sequence 18, Appl
42	17	1.2	990	10	US-09-974-300-523	Sequence 523, Ap
43	17	1.2	1041	10	US-09-778-844-129	Sequence 129, Ap
44	17	1.2	1043	10	US-09-778-844-130	Sequence 130, Ap
45	17	1.2	1048	10	US-09-822-849A-117	Sequence 117, Ap

#### ALIGNMENTS

RESULT 1  
US-09-908-988B-1  
Sequence 1, Application US/09908988B  
Patent No. US20020127690A1  
GENERAL INFORMATION:  
APPLICANT: OLSON, ERIC  
APPLICANT: SPENCER, JEFFREY A.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES  
FILE REFERENCE: MYOG:02805  
CURRENT APPLICATION NUMBER: US/09/908,988B  
CURRENT FILING DATE: 2000-07-18  
PRIOR APPLICATION NUMBER: 60/219,020  
PRIOR FILING DATE: 2000-07-18  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 1431  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (199)..(1296)  
US-09-908-988B-1  
Query Match 100.0%; Score 1431; DB 10; Length 1431;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAGCAGTGTACAGAGTGTCTGGAATATAGACAGGGTGAAGAGAGCTGTAGGGAGG 60  
1 AAGCAGTGTACAGAGTGTCTGGAATATAGACAGGGTGAAGAGAGCTGTAGGGAGG 60  
DB 1 AAGCAGTGTACAGAGTGTCTGGAATATAGACAGGGTGAAGAGAGCTGTAGGGAGG 60  
QY 61 GACAGACTCTTCCACAGAGGAGCAATAGCCGGATCCCAAGAAATCCAGTACGCTAAAC 120  
61 GACAGACTCTTCCACAGAGGAGCAATAGCCGGATCCCAAGAAATCCAGTACGCTAAAC 120  
DB 61 GACAGACTCTTCCACAGAGGAGCAATAGCCGGATCCCAAGAAATCCAGTACGCTAAAC 120  
QY 121 TGACCGAGGAAGGTGTCACAGGAGGAGGAGCAACGACAGGCGGACGAGGAGGAG 180  
121 TGACCGAGGAAGGTGTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
DB 121 TGACCGAGGAAGGTGTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
QY 181 GCTCCAGAGGGCCGGGAGTGAAGTCAAGGCTGTTCAAGCGGCTGCTAGGGAGTCG 240  
181 GCTCCAGAGGGCCGGGAGTGAAGTCAAGGCTGTTCAAGCGGCTGCTAGGGAGTCG 240

```

Db 181 GCTCCAGAGCGCGCGGATGAACTTACGCTGGGTTTCAAGCCGCTGCTAGGGGATCG 240
Qy 241 CACAACATGGACAACTTGGAGAGCAGCTATTTGCCCATCTCTCGGAGATGTTCTCC 300
Db 241 CACAACATGGACAACTTGGAGAGCAGCTATTTGCCCATCTCTCGGAGATGTTCTCC 300
Qy 301 AAGCCCGTGTATCTTGGCCCTGCCAACAACACCTGTGGCCGCAAGTGTGCCAAGAGCTC 360
Db 301 AAGCCCGTGTATCTTGGCCCTGCCAACAACACCTGTGGCCGCAAGTGTGCCAAGAGCTC 360
Qy 361 TTCCAGGCTCTTAATCTCTGTGGCAATCCGGGGCTCCACAAGCTGTCTTTCAGAGGA 420
Db 361 TTCCAGGCTCTTAATCTCTGTGGCAATCCGGGGCTCCACAAGCTGTCTTTCAGAGGA 420
Qy 421 CTTTCCGATGCCATCTTGTAGGCGAGGTTGCTCTGACAGGCTGTGTCTATGCC 480
Db 421 CTTTCCGATGCCATCTTGTAGGCGAGGTTGCTCTGACAGGCTGTGTCTATGCC 480
Qy 481 CTGACGCGAACCTGTAGTGGAGAACATCATGACATCTACAGCAGAGATCCTCCCG 540
Db 481 CTGACGCGAACCTGTAGTGGAGAACATCATGACATCTACAGCAGAGATCCTCCCG 540
Qy 541 CCACGACGCGCAAGGCTTAACAGCACTCATGTGTGAGAGCAGAGAGAGATC 600
Db 541 CCACGACGCGCAAGGCTTAACAGCACTCATGTGTGAGAGCAGAGAGATC 600
Qy 601 AACATCTACTGCTGAGCTGCGAGGAGTGCACCTGCTCTCTGCAAGGTTTCCGGCC 660
Db 601 AACATCTACTGCTGAGCTGCGAGGAGTGCACCTGCTCTCTGCAAGGTTTCCGGCC 660
Qy 661 CACAAGACTGTGAGAGTGGCCCTCTGCCCACATTACAAAGCCGCAAGAGTGAAGT 720
Db 661 CACAAGACTGTGAGAGTGGCCCTCTGCCCACATTACAAAGCCGCAAGAGTGAAGT 720
Qy 721 AGCATGCGATGCGATGCTGTGGGCGGCAATGACCGTGTGAGGAGTGAATCACCAG 780
Db 721 AGCATGCGATGCGATGCTGTGGGCGGCAATGACCGTGTGAGGAGTGAATCACCAG 780
Qy 781 ATGAGAGAGTGTGCGACCAATTGAGAGCAACAGCCGAGAGAGAGTGAATCACCAG 840
Db 781 ATGAGAGAGTGTGCGACCAATTGAGAGCAACAGCCGAGAGAGAGTGAATCACCAG 840
Qy 841 CAGAGGTTGAGACCGTGTGGGCGGTTTGGAGAGCGCAAGGCGGACTGCTTCAAGCA 900
Db 841 CAGAGGTTGAGACCGTGTGGGCGGTTTGGAGAGCGCAAGGCGGACTGCTTCAAGCA 900
Qy 901 CTGCCCCGCGAGAGAGAGAGTTCAGAGCGCTGCGGGGCTCATCCGCGAGTACGGA 960
Db 901 CTGCCCCGCGAGAGAGAGAGTTCAGAGCGCTGCGGGGCTCATCCGCGAGTACGGA 960
Qy 961 GACCACTTGGAGGCTCTCTCAAAAGTGTGAGTCCGCAATCCAGTCAATGAGAGAGCG 1020
Db 961 GACCACTTGGAGGCTCTCTCAAAAGTGTGAGTCCGCAATCCAGTCAATGAGAGAGCG 1020
Qy 1021 CAGATGCTCTTACCTTCCAGAGAGAGAGTGTGAGTCAACAAAGTGTGGGCAATGTG 1080
Db 1021 CAGATGCTCTTACCTTCCAGAGAGAGAGTGTGAGTCAACAAAGTGTGGGCAATGTG 1080
Qy 1081 AAGTGGAGCTGCGAGAGAGCGCGAGCGAGCTATGAGAGAGTGAAGCAATTCCTGTG 1140
Db 1081 AAGTGGAGCTGCGAGAGAGCGCGAGCGAGCTATGAGAGAGTGAAGCAATTCCTGTG 1140
Qy 1141 AGCGTGGAGAGCGGCGGCAAAATTTGCGAATTCGACTTCCAGCGCGCGCTGGG 1200
Db 1141 AGCGTGGAGAGCGGCGGCAAAATTTGCGAATTCGACTTCCAGCGCGCGCTGGG 1200
Qy 1201 GATGAAGAGATGACGACAGAGCTTTGATGAGAGAGAGAGGAGTGGGCTGAGAGAG 1260
Db 1201 GATGAAGAGATGACGACAGAGCTTTGATGAGAGAGAGAGGAGTGGGCTGAGAGAG 1260
Qy 1261 GAGCGCTGAGCTGCCAGAGAGCTGAGGCTGACATGACCGGACTGTGATCCAGAGCGC 1320
Db 1261 GAGCGCTGAGCTGCCAGAGAGCTGAGGCTGACATGACCGGACTGTGATCCAGAGCGC 1320

```

```

Db 1261 GAGCGCTGAGCTGCCAGAGAGCTGAGCGCTGACCTGACCCGACTCTGATCCAGAGCG 1320
Qy 1321 ACACCCGAAGCGGAGAGCCAAAGGATGCTGAGATGTGCGAGAGACACCGCGCCACCA 1380
Db 1321 ACACCCGAAGCGGAGAGCCAAAGGATGCTGAGATGTGCGAGAGACACCGCGCCACCA 1380
Qy 1381 GCTGCGCTTCCCGCCCGGGAAGGTTTCAATTAAGAGACTCAAGTGTCC 1431
Db 1381 GCTGCGCTTCCCGCCCGGGAAGGTTTCAATTAAGAGACTCAAGTGTCC 1431

```

## RESULT 2

```

US-09-908-988b-5
Sequence 5, Application US/09908988B
Patent No. US20020127690A1
GENERAL INFORMATION:
APPLICANT: OLSON, ERIC
APPLICANT: SPENCER, JEFFREY A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES
FILE REFERENCE: MYOG:0280S
CURRENT APPLICATION NUMBER: US/09/908,988B
CURRENT FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/219,020
PRIOR FILING DATE: 2000-07-18
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 1597
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (299)..(1327)
US-09-908-988b-5

```

```

Query Match 2.0%; Score 28; DB 10; Length 1597;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 505 AACATCATTTGACATCTCAAGCAGAGAGT 532
Db 596 AACATCATTTGACATCTCAAGCAGAGAGT 623

```

## RESULT 3

```

US-09-764-864-493
Sequence 493, Application US/09764864
Patent No. US20020132753A1
GENERAL INFORMATION:
APPLICANT: ROSEN ET AL.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT223
CURRENT APPLICATION NUMBER: US/09/764,864
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 493
LENGTH: 587
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-864-493

```

```

Query Match 1.9%; Score 27; DB 10; Length 587;
Best Local Similarity 100.0%; Pred. No. 0.00052;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 508 ATCATTTGACATCTCAAGCAGAGAGTCC 534
Db 445 ATCATTTGACATCTCAAGCAGAGAGTCC 471

```

1/31/00  
1/5/01



RESULT 4  
US-09-764-864-34  
; Sequence 34, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT723  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PAlM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 34  
; LENGTH: 1762  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-864-34

Query Match 1.9%; Score 27; DB 10; Length 1762;  
Best Local Similarity 100.0%; Pred. No. 0.00052;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 508 ATCATTGACATCTACAGCAGAGTCC 534  
|||||  
Db 467 ATCATTGACATCTACAGCAGAGTCC 493

RESULT 5  
US-09-917-800A-81/C  
; Sequence 81, Application US/09917800A  
; Patent No. US20020119462A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendick, Donna  
; APPLICANT: Porter, Mark  
; APPLICANT: Johnson, Kory  
; APPLICANT: Castle, Arthur  
; APPLICANT: Elashoff, Michael  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Molecular Toxicology Modeling  
; FILE REFERENCE: 44921-5038-US  
; CURRENT APPLICATION NUMBER: US/09/917,800A  
; CURRENT FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,040  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: US 60/222,880  
; PRIOR FILING DATE: 2000-11-02  
; PRIOR APPLICATION NUMBER: US 60/290,029  
; PRIOR FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: US 60/290,645  
; PRIOR FILING DATE: 2001-05-15  
; PRIOR APPLICATION NUMBER: US 60/292,336  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: US 60/295,798  
; PRIOR FILING DATE: 2001-06-06  
; PRIOR APPLICATION NUMBER: US 60/297,457  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,884  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/303,459  
; PRIOR FILING DATE: 2001-07-09  
; NUMBER OF SEQ ID NOS: 1740  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 81  
; LENGTH: 488  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AA818770  
US-09-917-800A-81

Query Match 1.3%; Score 19; DB 10; Length 488;  
Best Local Similarity 100.0%; Pred. No. 6.2;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1197 TGGGATCAAGAGTATGAC 1215  
|||||  
Db 390 TGGGATCAAGAGATGAC 372

RESULT 6  
US-09-822-849A-268  
; Sequence 268, Application US/09822849A  
; Patent No. US20020045170A1  
; GENERAL INFORMATION:  
; APPLICANT: Wong, Gordon G.  
; APPLICANT: Clark, Hilary  
; APPLICANT: Fechtel, Kim  
; APPLICANT: Agostino, Michael J.  
; APPLICANT: Howes, Steven H.  
; APPLICANT: Resnick, Richard J.  
; APPLICANT: Gulikota, Kamalakkar  
; APPLICANT: Graham, James R.  
; APPLICANT: Genetics Institute, Inc.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS  
; FILE REFERENCE: GIN 6403  
; CURRENT APPLICATION NUMBER: US/09/822,849A  
; CURRENT FILING DATE: 2001-09-04  
; PRIOR APPLICATION NUMBER: 60/195,582  
; PRIOR FILING DATE: 2000-04-06  
; NUMBER OF SEQ ID NOS: 598  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 268  
; LENGTH: 2513  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-822-849A-268

Query Match 1.3%; Score 19; DB 10; Length 2513;  
Best Local Similarity 100.0%; Pred. No. 6.3;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 137 CACAGGCGAGGGGAGAGGC 155  
|||||  
Db 1595 CACAGGCGAGGGGAGAGGC 1613

RESULT 7  
US-09-960-352-207  
; Sequence 207, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960,352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 207  
; LENGTH: 399  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (359)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: 01-LIB34-072-01-E1-A1  
US-09-960-352-207

Query Match 1.3%; Score 18; DB 10; Length 399;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	795	CCAGACCATTTGAGGACAA	812
Db	275	CCAGACCATTTGAGGACAA	292

## RESULT 8

```

US-09-960-352-13152
? Sequence 13152, Application US/09960352
? Patent No. US20020137139A1
? GENERAL INFORMATION:
? APPLICANT: Warren, Wesley C.
? APPLICANT: Tso, Nengbing
? APPLICANT: Byatt, John C.
? APPLICANT: Mathalagan, Nagappan
? TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
? FILE REFERENCE: 16511.006/37-21(10298)C
? CURRENT APPLICATION NUMBER: US/09/960,352
? CURRENT FILING DATE: 2001-09-24
? NUMBER OF SEQ. ID NOS: 15112
? SEQ. ID NO 13152
? LENGTH: 444
? TYPE: DNA
? ORGANISM: Bos taurus
? OTHER INFORMATION: Clone ID: 56-LIB34-001-Q1-E1-F8
US-09-960-352-13152

```

Query Match	1.3%	Score 18	DB 10	Length 444
Best Local Similarity	100.0%	Pred. No. 20		
Matches 18	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	795	CCAGACCATTGAGGACAA	812
Db	266	CCAGACCATTGAGGACAA	283

RESULT 9  
US-09-917-800A-577

Sequence 577, Application US/09917800A  
Patent No US00702119462A1  
GENERAL INFORMATION:  
APPLICANT: Mendrick, Donna  
APPLICANT: Porter, Mark  
APPLICANT: Johnson, Kory  
APPLICANT: Castle, Arthur  
APPLICANT: Elashoff, Michael  
APPLICANT: Gene Logic, Inc.  
TITLE OF INVENTION: Molecular Toxicology Modeling  
FILE REFERENCE: 44921-5038-US  
CURRENT APPLICATION NUMBER: US/09/917,800A  
PRIORITY FILING DATE: 2001-07-31  
PRIORITY APPLICATION NUMBER: US 60/222,040  
PRIORITY FILING DATE: 2000-07-31  
PRIORITY APPLICATION NUMBER: US 60/222,880  
PRIORITY FILING DATE: 2000-11-02  
PRIORITY APPLICATION NUMBER: US 60/290,029  
PRIORITY FILING DATE: 2001-05-11  
PRIORITY APPLICATION NUMBER: US 60/290,645  
PRIORITY FILING DATE: 2001-05-15  
PRIORITY APPLICATION NUMBER: US 60/292,336  
PRIORITY FILING DATE: 2001-05-22  
PRIORITY APPLICATION NUMBER: US 60/295,798  
PRIORITY FILING DATE: 2001-06-06  
PRIORITY APPLICATION NUMBER: US 60/297,457  
PRIORITY FILING DATE: 2001-06-13  
PRIORITY APPLICATION NUMBER: US 60/298,884  
PRIORITY FILING DATE: 2001-06-19  
PRIORITY APPLICATION NUMBER: US 60/303,459  
PRIORITY FILING DATE: 2001-07-09  
NUMBER OF SEQ ID NOS: 1740  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 577

```

; LENGTH:445
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 A1028973
US-09-917-800A-577

```

Query Match	1.3%;	Score 18;	DB 10;	Length 445;
Best Local Similarity	100.0%;	Pred. No. 20;		
Matches 18; Conservative	0;	Mismatches	0;	Gaps 0

QY	1377	CCAAGCTCGGCTTCCGC	1394
Db	212	CCAAGCTCGGCTTCCGC	229

RESULTS 10  
ITS-09-736

```

1 Sequence 753, Application US/09736457
2 Patent No. US20020168637A1
3 GENERAL INFORMATION:
4 APPLICANT: Wang, Tongtong
5 APPLICANT: Bangur, Chaitanya S.
6 APPLICANT: Iodes, Michael A.
7 APPLICANT: Fanger, Gary
8 APPLICANT: Vedvick, Tom
9 APPLICANT: Carter, Darrick
10 APPLICANT: Retter, Marc
11 APPLICANT: Mannion, Jane
12 APPLICANT: Pan, Liqun
13 APPLICANT: Wang, Ailun
14 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
15 TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
16 FILE REFERENCE: 210121.478615
17 CURRENT APPLICATION NUMBER: US/09/736,457
18 CURRENT FILING DATE: 2000-12-13
19 NUMBER OF SEQ. ID NOS: 1864
20 SOFTWARE: FastSeq for Windows Version 3.0
21 SEQ. ID NO 753
22 LENGTH: 467
23 TYPE: DNA
24 ORGANISM: Homo sapien
25 FEATURE:
26 NAME/KEY: misc_feature
27 LOCATION: (1)..(467)
28 OTHER INFORMATION: n = A,T,C or G
29 -S-09-736-457-753

```

Query Match	1.3%	Score 18:	DB 9;	Length 467;
Best Local Similarity	100.0%	Pred. No. 20;		
Matches 18;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	636	CTCTCTCTGCAAGGTTT	653
Db	442	CTCTCTCTGCAAGGTTT	425

RESULT 11

US-95-002-941, 753/C  
Sequence 753, Application US-099020944  
Patent No. US20020172552A1  
GENERAL INFORMATION:  
APPLICANT: Henderson, Robert A.  
APPLICANT: Wang, Tongtong  
APPLICANT: Watanabe, Yoshihiro  
APPLICANT: Johnson, Jeffrey C.  
APPLICANT: Retter, Marc W.  
APPLICANT: Mernerakus, Margarita  
APPLICANT: Carter, Darrick  
APPLICANT: Fanger, Gary R.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Bangur, Chaltanya S.  
APPLICANT: McNab, Andria

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902.941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 753
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1, 15, 77, 314, 317, 335, 419
; OTHER INFORMATION: n = A,T,C or G
US-09-902-941-753

Query Match          1.3%; Score 18; DB 9; Length 467;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 CTCTCTGCAAGGTTT 653
    |||
Db 442 CTCTCTGCAAGGTTT 425

RESULT 12
US-09-822-849A-580
; Sequence 580, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 580
; LENGTH: 746
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-580

Query Match          1.3%; Score 18; DB 10; Length 746;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 904 GCCCGGAGCAGAGAG 921
    |||
Db 405 GCCCGGAGCAGAGAG 422

RESULT 13
US-09-833-381-303/C
; Sequence 303, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
```

```

; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 303
; LENGTH: 860
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)...(860)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-303

Query Match          1.3%; Score 18; DB 10; Length 860;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 275 GCCCATCTGCTGGAGA 292
    |||
Db 217 GCCCATCTGCTGGAGA 200

RESULT 14
US-10-044-090-508
; Sequence 508, Application US/10044090
; Patent No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 508
; LENGTH: 3033
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 082155CH1
US-10-044-090-508

Query Match          1.3%; Score 18; DB 12; Length 3033;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 274 TGCCCCATCTGCTGGAG 291
    |||
Db 355 TGCCCCATCTGCTGGAG 372

RESULT 15
US-09-917-800A-1593
; Sequence 1593, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castile, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
```

PRIOR FILING DATE: 2001-05-11  
PRIOR APPLICATION NUMBER: US 60/290,645  
PRIOR FILING DATE: 2001-05-15  
PRIOR APPLICATION NUMBER: US 60/292,336  
PRIOR FILING DATE: 2001-05-22  
PRIOR APPLICATION NUMBER: US 60/295,798  
PRIOR FILING DATE: 2001-06-06  
PRIOR APPLICATION NUMBER: US 60/297,457  
PRIOR FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/298,884  
PRIOR FILING DATE: 2001-06-19  
PRIOR APPLICATION NUMBER: US 60/303,459  
NUMBER OF SEQ ID NOS: 1740  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 1593  
LENGTH: 4153  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM\_012999  
US-09-917-800A-1593

Query Match 1.3%; Score 18; DB 10; Length 4153;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 608 ACTGCTGAGCTGCGAGG 625  
|||||  
DB 2643 ACTGCTGAGCTGCGAGG 2660

RESULT 16  
US-09-263-959-1/c  
Sequence 1, Application US/09263959  
Patent No. US20020150891A1  
GENERAL INFORMATION:  
APPLICANT: Hood, Leroy E.  
APPLICANT: Rowen, Lee  
TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTI  
NUMBER OF SEQUENCES: 1279  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/263,959  
FILING DATE: 05-MAR-1999  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mcmasters, David D.  
REGISTRATION NUMBER: 33,963  
REFERENCE/DOCKET NUMBER: 920010.426C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 684973 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-263-959-1

Query Match 1.3%; Score 18; DB 10; Length 684973;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1034 ACTTCACAGCAGCAAGG 1051  
|||||  
DB 526448 ACTTCACAGCAGCAAGG 526431

RESULT 17  
US-09-776-695-22  
Sequence 22, Application US/09776695  
Patent No. US20020068283A1  
GENERAL INFORMATION:  
APPLICANT: Boeke, Jef  
Brachmann, Rainer  
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING P53  
MUTATIONS WHICH SUPPRESS P53 CANCER MUTA- TIONS  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Wilcoff  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20001

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,695  
FILING DATE: 26-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/795,006  
FILING DATE: 2001-03-28  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A  
REGISTRATION NUMBER: 32141  
REFERENCE/DOCKET NUMBER: 01107.03170  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
TELEX: <unknown>  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 133 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-776-695-22

Query Match 1.2%; Score 17; DB 10; Length 133;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 CCATGAGAGCCCGCAG 1023  
|||||  
DB 42 CCATGAGAGCCCGCAG 58

RESULT 18  
US-09-878-574-9448/c  
Sequence 9448, Application US/09878574  
Patent No. US20020110548A1  
GENERAL INFORMATION:  
APPLICANT: Byrum, Joseph R.  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Thompson, Michael D.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

```

; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 9448
; LENGTH: 192
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701102502H1
US-09-878-574-9448

Query Match          1.2%; Score 17; DB 10; Length 192;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1248 GGGGCTGGAGAGAGAGC 1264
      |||||||||||||||
Db 162 GGGGCTGGAGAGAGAGC 146

RESULT 19
US-09-864-761-22504
; Sequence 22504, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmice-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
```

```

; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22504
; LENGTH: 230
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004477.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
; OTHER INFORMATION: NT HIT: X84060.1, EVALUE 1.00e-127
; OTHER INFORMATION: SWISSPROT HIT: Q14494, EVALUE 1.00e-38
; OTHER INFORMATION: EST_HUMAN HIT: BE897190.1, EVALUE 1.00e-116
US-09-864-761-22504
```

```
Query Match          1.2%; Score 17; DB 10; Length 230;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 487 CGAAGCTGCTAGTGA 503
      |||||||||||||||
Db 166 CGAAGCTGCTAGTGA 182
```

```

RESULT 20
US-09-923-876-2132
; Sequence 2132, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Laigudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 2132
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700160095H1
; LOCATION: 190, 220
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-2132
```

```
Query Match          1.2%; Score 17; DB 10; Length 231;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 499 GTGAGACATCATGA 515
      |||||||||||||||
Db 123 GTGAGACATCATGA 139
```

Search completed: December 3, 2002, 12:14:06  
Job time : 535 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 06:49:10 ; Search time 2791 Seconds

(without alignments)  
14921.575 Million cell updates/sec

Title: us-09-908-988b-1

Sequence: 1 aagagagctagacagagatgtc.....ataagagactcaaggtcc 1431

Scoring table: IDENTITY-NUC  
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402678 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: 1  
1: gb\_pat\*  
2: gb\_hvg\*  
3: gb\_in\*  
4: gb\_ov\*  
5: gb\_ov\*  
6: gb\_pat\*  
7: gb\_ph\*  
8: gb\_pl\*  
9: gb\_pl\*  
10: gb\_ro\*  
11: gb\_sts\*  
12: gb\_sy\*  
13: gb\_un\*  
14: gb\_vl\*  
15: em\_ba\*  
16: em\_fun\*  
17: em\_hum\*  
18: em\_in\*  
19: em\_mu\*  
20: em\_om\*  
21: em\_or\*  
22: em\_ov\*  
23: em\_pal\*  
24: em\_ph\*  
25: em\_pl\*  
26: em\_ro\*  
27: em\_sts\*  
28: em\_un\*  
29: em\_vl\*  
30: em\_hvg\_hum\*  
31: em\_hvg\_inv\*  
32: em\_hvg\_other\*  
33: em\_hvg\_mus\*  
34: em\_hvg\_pln\*  
35: em\_hvg\_rod\*  
36: em\_hvg\_man\*  
37: em\_hvg\_vtl\*  
38: em\_sy\*  
39: em\_higo\_hum\*  
40: em\_higo\_mus\*  
41: em\_higo\_other\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1431	100.0	1431	6	AX418848
2	1431	100.0	1448	10	AF294790
3	772.4	54.0	1329	6	HSR291714
4	626.4	43.8	1500	6	AX060632
5	484.6	33.9	1597	6	AX418852
6	483.6	33.8	2097	9	HSR291713
7	482	33.7	1756	9	HSR276484
8	482	33.7	1764	6	AX056942
9	482	33.7	1764	9	AX056942
10	482	33.7	1791	9	AF353673
11	476.8	33.3	1861	10	AY059627
12	475.6	33.2	1211	9	BC015717
13	450.8	31.5	2590	6	AX418850
14	450.6	31.5	1750	9	BC007750
15	450.6	31.5	1925	9	AK091728
16	450.6	31.5	2202	9	HSR291712
17	450.6	31.5	2634	9	AK091310
18	447.4	31.3	1810	9	HSR243488
19	447.4	31.3	2098	9	HSR243489
20	367	25.6	209211	2	AC084883
21	350	24.5	1746	9	AB047601
22	267.4	18.7	531	6	AX313116
23	248.6	17.4	1903	9	AF361946
24	237.8	16.6	199296	2	AC114619
25	189.4	13.2	118847	2	AC013413
26	173.4	12.1	49743	2	AC114541
27	162.6	11.4	164772	2	AC095184
28	154	10.8	123280	2	AC025562
29	107.4	7.5	121721	9	AC090195
30	107.4	7.5	183912	2	AC021240
31	107.2	7.5	168672	2	AC128256
32	107.2	7.5	181514	2	AC095643
33	106.8	7.5	176006	9	AL391650
34	106.4	7.4	20092	10	AF627253
35	106.4	7.4	73989	2	AC101573
36	106.4	7.4	196461	10	AL627314
37	105.8	7.4	170980	2	AF277371
38	104.4	7.3	174303	2	AC099104
39	103.8	7.3	160040	2	AC105674
40	103	7.2	174303	2	AC099104
41	100.2	7.0	160040	2	AC105674
42	97	6.8	1875	9	AK026882
43	95.4	6.7	3127	9	AK096158
44	94.6	6.6	168672	2	AC128256
45	90.6	6.3	2716	9	AB056399

## ALIGNMENTS

RESULT 1  
LOCUS AX418848 1431 bp DNA linear PAT 18-JUN-2002  
DEFINITION Sequence 1 from Patent WO0206318.  
ACCESSION AX418848  
VERSION AX418848.1 GI:21523712  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Olson, E.N. and Spencer, J.A.  
TITLE Methods and compositions for stabilizing microtubules and intermediate filaments in striated muscle cells

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0206318-A 1 24-JAN-2002;  
BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)  
FEATURES Location/Qualifiers  
source 1. 1431  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
CDS 199..1299  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD35455.1"  
/db\_xref="GI:21523713"

BASE COUNT 338 a 384 c 473 g 236 t  
ORIGIN

Query Match 100.0%; Score 1431; DB 6; Length 1431;  
Best Local Similarity 100.0%; Pred. No. 1.1e-267;  
Matches 1431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGAGTGTAGACAGAGTGTCTGGAATAGACAGAGGCTGAGAGAGCTTTAGGGAGAG 60  
DB 1 AAGAGTGTAGACAGAGTGTCTGGAATAGACAGAGGCTGAGAGAGCTTTAGGGAGAG 60  
QY 61 GACAGAGCTCTTCCAAAGAGGAGCAATAGCCGGGATCCCAAGATCCAGTCAACCTTAAC 120  
DB 61 GACAGAGCTCTTCCAAAGAGGAGCAATAGCCGGGATCCCAAGATCCAGTCAACCTTAAC 120  
QY 121 TGACCGAGAGAGGCTGACAGAGGAGGAGGAGCAAGAGAGGAGGAGGAGGAGGAG 180  
DB 121 TGACCGAGAGAGGCTGACAGAGGAGGAGGAGGAGCAAGAGAGGAGGAGGAGGAG 180  
QY 181 GCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
DB 181 GCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
QY 241 CACAACATGACAACTTGAGAGAGAGCTCAATTTGCCCATCTGCTGGAGATGTTCTCC 300  
DB 241 CACAACATGACAACTTGAGAGAGAGCTCAATTTGCCCATCTGCTGGAGATGTTCTCC 300  
QY 301 AAGCCGCTGATCTTCCCTGCCAACACAACCTGTGCCGCAAGTGTGCCAACGAGTC 360  
DB 301 AAGCCGCTGATCTTCCCTGCCAACACAACCTGTGCCGCAAGTGTGCCAACGAGTC 360  
QY 361 TTCAGGCTCTATTCCTGTGGCAATCCCGGGGCTCCACAACCGTGTCTTCCAGAGGA 420  
DB 361 TTCAGGCTCTATTCCTGTGGCAATCCCGGGGCTCCACAACCGTGTCTTCCAGAGGA 420  
QY 421 CGTTTCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
DB 421 CGTTTCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480  
QY 481 CTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
DB 481 CTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
QY 541 CCACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
DB 541 CCACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
QY 601 AACATCTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
DB 601 AACATCTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
QY 661 CACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
DB 661 CACAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720

QY 721 ACCATGCGATTCGCGATGCTGTGGCGGCAATGACCGTGTGACAGGAGCTGATCACCAG 780  
DB 721 AGCGATGCGATTCGCGATGCTGTGGCGGCAATGACCGTGTGACAGGAGCTGATCACCAG 780  
QY 781 ATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840  
DB 781 ATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840  
QY 841 CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
DB 841 CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
QY 901 CTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
DB 901 CTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
QY 961 GACCACTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020  
DB 961 GACCACTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020  
QY 1021 CAGATGCTCTCTACCTTCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
DB 1021 CAGATGCTCTCTACCTTCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080  
QY 1081 AAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140  
DB 1081 AAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140  
QY 1141 AGCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
DB 1141 AGCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
QY 1201 GATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
DB 1201 GATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
QY 1261 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320  
DB 1261 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320  
QY 1321 ACACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380  
DB 1321 ACACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380  
QY 1381 GCTCGGCTTCCCGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1431  
DB 1381 GCTCGGCTTCCCGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1431

RESULT 2  
AF294790 1448 bp mRNA linear ROD 08-SEP-2000  
LOCUS AF294790  
DEFINITION Mus musculus RING-finger protein MURF mRNA, complete cds.  
ACCESSION AF294790  
VERSION AF294790.1 GI:9945009  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1448)  
Spencer, J.A., Eliazar, S., Ilaria, R.L. Jr., Richardson, J.A. and  
Olson, E.N.  
TITLE  
Regulation of microtubule dynamics and myogenic differentiation by  
MURF, a striated muscle RING-finger protein  
JOURNAL  
MOL. CELL. BIOL. 20(4):771-784 (2000)  
MEDLINE  
20011220  
PUBMED  
10953002  
REFERENCE  
2 (bases 1 to 1448)  
Spencer, J.A. and Olson, E.N.  
AUTHORS  
Direct Submission  
TITLE  
Submitted (09-AUG-2000) Molecular Biology, UTSW, 6000 Harry Hines  
Blvd., Dallas, TX 75390-9148, USA



FEATURES	Location/Qualifiers
source	1. .1448

CDS

BASE COUNT	355 a	384 c	473 g	236 t
ORIGIN				

Query Match	100.0%;	Score 1431;	DB 10;	Length 1448;
Best Local Similarity	100.0%;	Pred. No. 1.1e-267;		
Matches 1431, Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Db	1	AAGGAGTGTAGACGAGCTGTCTGGAAATACACAGGGGTGAGAGACTGTTAGGGGAAG	60
OY	61	GACAGGACTCTTCCAAAGAGGAGCAATAGCCGGGATTCCCAAGAAATCCAGTACACCTTAAC	120
Db	61	GACAGGACTCTTCCAAAGAGGAGCAATAGCCGGGATTCCCAAGAAATCCAGTACACCTTAAC	120
OY	121	TGACCGAGGAAGGGTGTGCACAGGCGAGGGGAGAAAGGCCAACACACAGGGCCACAGCGAGCAG	180
Db	121	TGACCGAGGAAGGGTGTGCACAGGCGAGGGGAGAAAGGCCAACACACAGGGCCACAGCGAGCAG	180
OY	181	GCTCCAGAGCCGCGGGGATGAATTCACGGTGGGTTTCAAGCCGCTGCTAGGGATGCG	240
Db	181	GCTCCAGAGCCGCGGGGATGAATTCACGGTGGGTTTCAAGCCGCTGCTAGGGATGCG	240
OY	241	CACAAATGSGAACATTGGGAAGACACTATTTGGCCCATCTGCGCGGAGATGTCTCC	300
Db	241	CACAAATGSGAACATTGGGAAGACACTATTTGGCCCATCTGCGCGGAGATGTCTCC	300
OY	301	AAGCCCTGTGTATCTTGCCCTGCGCAACAGACCTGTGCGCAAGTGTGCAACGAGCTC	360
Db	301	AAGCCCTGTGTATCTTGCCCTGCGCAACAGACCTGTGCGCAAGTGTGCAACGAGCTC	360
OY	361	TTTCAGGSCCTTAATCCTCTGTGGCAATCCCGGGGCTCCACAAACGATGTCTTACGAGGA	420
Db	361	TTTCAGGSCCTTAATCCTCTGTGGCAATCCCGGGGCTCCACAAACGATGTCTTACGAGGA	420
OY	421	CGTTCCGATGCCATCTGTTAGGACAGAGGTTTCCGTGGACAGGATGGTGCTATAGG	480
Db	421	CGTTCCGATGCCATCTGTTAGGACAGAGGTTTCCGTGGACAGGATGGTGCTATAGG	480
OY	481	CTGCAGCGGAACCTGTAGTGAGAGAACATCTATTACATCTTACAAAGCAGGAGTCCCTCCGG	540
Db	481	CTGCAGCGGAACCTGTAGTGAGAGAACATCTATTACATCTTACAAAGCAGGAGTCCCTCCGG	540
OY	541	CCACTGTCAGCGCCAAAGGCTGAACAGACCTCATGTGTGAGAGACGAGAGAGAGATC	600
Db	541	CCACTGTCAGCGCCAAAGGCTGAACAGACCTCATGTGTGAGAGACGAGAGAGAGATC	600
OY	601	AACATCTACTGCTAGTGTGGAGAGGTGCCACAGCTCTCTGTCAGAGTTTGTGGGCG	660
Db	601	AACATCTACTGCTAGTGTGGAGAGGTGCCACAGCTCTCTGTCAGAGTTTGTGGGCG	660
OY	661	CACAGGAGCTGTGAGAGTGGCCCTCTGGCCACCAATTTACAAGCCCAAGAGAGTGA	720
Db	661	CACAGGAGCTGTGAGAGTGGCCCTCTGGCCACCAATTTACAAGCCCAAGAGAGTGA	720
OY	721	AGCGATGGCAGTGCATCTGTGTGGCGGGCAATGACCTGTGTGACGACGATGATACCCAG	780

Db	721	AGCGATGGCATTGGGATGCTGTGGGGCAATACCGGTGTGCAGGCAATGATCACCAG	780
QY	781	ATGAGAGAGGTGTGCCAGACCATTTAGAGACACACAGCCGACAGACGAAGCAACTGTTAAAC	840
Db	781	ATGAGAGAGGTGTGCCAGACCATTTAGAGACACACAGCCGACAGACGAAGCAACTGTTAAAC	840
QY	841	CAGAGGTTCCGAGACCCTGTGCGGTTTTTGGAGAGGCGCAAGGGGCGAAACTGCTTCAAGCA	900
Db	841	CAGAGGTTCCGAGACCCTGTGCGGTTTTTGGAGAGGCGCAAGGGGCGAAACTGCTTCAAGCA	900
QY	901	CTGGCCCGGAGACAGAGAGAGAGAGTTGCAGCGCGTGGCGGGGCTTCATCCGCCAGTACGGGA	960
Db	901	CTGGCCCGGAGAGAGAGAGAGAGTTGCAGCGCGTGGCGGGGCTTCATCCGCCAGTACGGGA	960
QY	961	GACCACTTGGAGGGCTCCTCAAGATGGTGGTAGTCCGCAATCCAGTCCATGAGAGCGG	1020
Db	961	GACCACTTGGAGGGCTCCTCAAGATGGTGGTAGTCCGCAATCCAGTCCATGAGAGCGG	1020
QY	1021	CAGATGCTCTTACCTTCACCTTCAGACAGGCAAAAGAGAGCTGATCAACAAAGTGGGGCAATGTGC	1080
Db	1021	CAGATGCTCTTACCTTCACCTTCAGACAGGCAAAAGAGAGCTGATCAACAAAGTGGGGCAATGTGC	1080
QY	1081	AAGGTGAGCTGTGGCAGACAGCGCCGAGGCGTATGTAGAGAGATGAGACATTTCTGTG	1140
Db	1081	AAGGTGAGCTGTGGCAGACAGCGCCGAGGCGTATGTAGAGAGATGAGACATTTCTGTG	1140
QY	1141	AGCGTGGAGCAGTGTGGCGGAATGTGGCAACCATTCAGCTTCACGCGGGCGCGCTGGG	1200
Db	1141	AGCGTGGAGCAGTGTGGCGGAATGTGGCAACCATTCAGCTTCACGCGGGCGCGCTGGG	1200
QY	1201	GATGAAGAGGATGACAGACATGGCTTTGGATGGGGAGAGAGGCAATGCGGGCTGGAGAG	1260
Db	1201	GATGAAGAGGATGACAGACATGGCTTTGGATGGGGAGAGAGGCAATGCGGGCTGGAGAG	1260
QY	1261	GAGCGGCTGGAGAGCTGCCAGGAAGGCTCAGGCGCTGCACCTGACCTTGATCCAGAGGCG	1320
Db	1261	GAGCGGCTGGAGAGCTGCCAGGAAGGCTCAGGCGCTGCACCTGACCTTGATCCAGAGGCG	1320
QY	1321	ACACCCGGAAGCGGAGCCAAAGGATGCTGAGAGATCTGGCGACAGACCAACCGGCGCCACAA	1380
Db	1321	ACACCCGGAAGCGGAGCCAAAGGATGCTGAGAGATCTGGCGACAGACCAACCGGCGCCACAA	1380
QY	1381	GCTGGGCTTCCCGCCCGGGGAAGGTTCATTAAGAAGCACTCAAGTCCC	1441
Db	1381	GCTGGGCTTCCCGCCCGGGGAAGGTTCATTAAGAAGCACTCAAGTCCC	1441

RESULT 3  
HSA291714 1329 bp mRNA linear PRI 06-JUL-2001

LOCUS HSA291714  
DEFINITION Homo sapiens mRNA for RNF30 gene for ring finger protein 30.  
ACJ291714  
VERSION AJ291714.1 GI:13160387  
KEYWORDS MURF-3 gene; muscle specific ring finger protein.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1329)  
Centner,T., Yano,J., Kimura,E., MELHINY,A.S., Pelin,K., Witt,C.C., Bang,M.L., Trombitas,K., Granzier,H., Gregorio,C.C., Sorimachi,H. and Iabetti,S.  
Identification of muscle specific ring finger proteins as potential regulators of the titin kinase domain  
J. Mol. Biol. 306 (4), 717-726 (2001)  
2 (bases 1 to 1329)  
Centner,T.  
Direct Submission  
Submitted (12-FEB-2001) Centner T., Structure and Biocomputing, EMBL-Heidelberg, Meyerhofstrasse 1, Heidelberg 69112, Germany



ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1500)  
AUTHORS Lal, P., Yue, H., Tang, Y. T., Baughn, M. R., Azimzal, Y. and Tran, B.  
TITLE Human transcriptional regulator proteins  
JOURNAL Patent: WO 0078954-A 54 28-DEC-2000;  
Incyte Genomics, Inc. (US)

FEATURES  
source 1..1500  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 401 a 378 c 439 g 282 t

ORIGIN

Query Match 43.8%; Score 626.4; DB 6; Length 1500;  
Best Local Similarity 87.5%; Pred. No. 1.4e-111;  
Matches 721; Conservative 0; Mismatches 96; Indels 7; Gaps 3;

QY 118 AACGACCGAGGAGGCTGCACAGGCGAG-GGGAGAGGCCAAGCAGCGGCCACAGCGAG 176  
DB 214 ATCTAAGCGAGGAGGCTGTACAGCGAGTGAAGGCCAGAGGCGGCCAG 273  
QY 177 GGAGGCTCCAGAGCGCGGGGGATGAACTTACGCTGGGTTTCAAGCGCGCTGTAAGGGA 236  
DB 274 GGACGACG-----ACCGAGGGATGAACCTTACAGTGGGTTTCAAGCGCGCTGTAAGGGA 328  
QY 237 TCGCGCAACATGAGCACTGTGAGAGAGCAGCTCATTTGCCCATCTGCTGAGATGTT 296  
DB 329 TGCACACACATGAGCAACCTGTGAGAGAGCAGCTCATCTGCCCCATCTGCGGAGATGTT 388  
QY 297 CTCGAAGCGCGTGTATCTTGGCCCTGCCAACAACCTGCGCGGAGTGGCAACGA 356  
DB 389 CTCGAAGCGAGTGTATCTTGGCCCTGCCAACAACCTGCGCGGAGTGGCAACGA 448  
QY 357 CGTCTTCCAGGCGCTTAACTCTGTGGCAATCCGGGGCTCCCAACAAGGTTCTTCAGG 416  
DB 449 CGTCTTCCAGGCGCTTAACTCTGTGGCAATCCGGGGCTCCCAACAAGGTTCTTCAGG 508  
QY 417 AGGAGCTTTCCATGCGCCATCTGTGAGGACAGAGTGTCTCTGAGACAGGCAATGGTGTCTA 476  
DB 509 AGGCGCTTTCCGCTGCTGAGGCTGTGAGGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 568  
QY 477 TGGCTGTGACGCGGACCTGTGTAGTGAAGACATGATGACATCTTACAAGCAGAGTCTCTC 536  
DB 569 CGGCTGTGACGCGGAGAACTGTGTAGGAGAACATATGACATTTACAAGCAGAGTCTCTC 628  
QY 537 CCGGCACTGCGAGCGCAAGGCTGAAACAGCACCTCATGTGTGAGAGCAGAGGAGGAGAA 596  
DB 629 CAGGCGGCTGCACTCCCAAGGCTGAGAGCACCTCATGTGTGAGAGGAGGAGGAGGAGAA 688  
QY 597 GATCAACATCTACTGCTGCTGTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 656  
DB 689 GATCAACATCTACTGCTGCTGTGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 748  
QY 657 CGCCCAAGAGGAGTGTAGTGTGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 716  
DB 749 TGGCCCAAGAGGAGTGTAGTGTGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 808  
QY 717 GCTGAGGAGGAGTGTGAGGCTGTGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 776  
DB 809 GCTGAGGAGGAGTGTGAGGCTGTGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 868  
QY 777 CCGAGGAGGAGGAGTGTGAGGCTGTGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 836  
DB 869 ACAGATGAGGAGGAGTGTGAGGCTGTGAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 928  
QY 837 AACACGAGGTTGAGAGCCTGTGTGCGCGGCTTTGGAGAGAGGCGGAGGAGGAGGAGGAGG 896  
DB 929 AACACGAGGTTGAGAGCCTGTGTGCGCGGCTTTGGAGAGAGGCGGAGGAGGAGGAGGAGG 988  
QY 897 AGCACTGGCGCGGAGGAGC-AGGAGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGG 939

DB 989 GCGCTGGCGCGGAGAACAGCAGGAGCTTCAACGCGGATCCG 1032

RESULT 5  
AX418852  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AX418852  
Sequence 5 from Patent WO0206318.  
AX418852  
AX418852.1 GI:21523716

house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1  
AUTHORS Olson, E.N. and Spencer, J.A.  
TITLE Methods and compositions for stabilizing microtubules and intermediate filaments in striated muscle cells  
JOURNAL Patent: WO 0206318-A 5 24-JAN-2002;  
BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)

FEATURES  
source 1..1597  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
299..1330  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD35457.1"  
/db\_xref="GI:21523717"  
/translation="MDYKSLTPDGNAMENLEKOLICPTCELEMTKPVVILPCOHNLK  
RKANDIRQANPWTNRGVSMSGGRRCPSCHREVIIMDHVIGLDRNLVENVIT  
DIYKQESRPLQKSHPMCKEHEDEKIIYGLTEVPYSLCKYGAHQAEVAPLQ  
SIFQKQELSNICISMLVANDRYOTIISLDSCVRKENSHEVSELSOKRDTLYA  
ILDEKSELRLITQEOEKIKGFIALLIYREOEKRSKIVETALIOSLDEFGATFL  
SSAKGLISYIYBASKGCOLKTEQGEPMNDYFTLDEHIAELRAIDFGTKGCDVTC  
LTFEORRS"

BASE COUNT 405 a 404 c 438 g 350 t

ORIGIN

Query Match 33.9%; Score 484.6; DB 6; Length 1597;  
Best Local Similarity 69.2%; Pred. No. 4.6e-84;  
Matches 677; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

QY 226 CTGCTAGGGGAGTGGCAACAACATGACACTTGGAGAGAGGCTCATTTGGCCCATCTGC 285  
DB 317 CTGATTCCTGATGAGAAAGCTATGAGAGAACCTGAGAGGAGTCTGCCCATCTGC 376  
QY 286 CTGGAATGTTCTTCCAAAGCCGCTGTGATCTTGGCCGCAACACACACTGTGGCGCAAG 345  
DB 377 CTGAGAGTGTTCACAAAGCCTGTGATCTTGGCCGCAACACACACTGTGGCGCAAG 436  
QY 346 TGTGCCAAGACGCTTCTCCAGGCGCTTAACTCTGTGGAATCCGCGGCTCCACAGC 405  
DB 437 TGTGCCAAGACATCTTCCAGGCGCTTAACTCTGTGGAATCCGCGGCTCCACAGC 496  
QY 406 GTGTCTTCCAGGAGAGCTTCCGATGCCCATTGTGTAGGACGAGGTTGTCTGGACAGG 465  
DB 497 TCCATGCTTGAAGGTGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 556  
QY 466 CATGCTGTATGAGGCGCTGAGCGGAGCCTGTGATGAGAGACATCATTTGACATCTACAG 525  
DB 557 CAGGCGGTGTACGCGCTGCGAGAGAACCTGTGTGGAAGAAACATCATTTGACATCTACAG 616  
QY 526 CAGAGTCTTCCGCGGCTGTGACGCGGAGGCTGAAACAGCAGCTCATGTGTGAGAGAGC 585  
DB 617 CAGAGTCTTCCGCGGCTGTGACGCGGAGGCTGAAACAGCAGCTCATGTGTGAGAGAGC 673  
QY 586 GAGGAGAGAGATCATCATCTACTGCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 645  
DB 674 GAGGAGAGAGATCATCATCTACTGCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 733  
QY 646 AAGTTTTCGCGCGCCCAAGGAGTGTGAGGTTGCGCCCTCTGCCCACATTTACAAGGC 705

```

Db 734 AAGGTGTTGGGGCTCACCAGGCGCTGTGAGTTGCCCTTTGGCAAGCATCTTCCAAAGA 793
Qy 706 CAGAAAGTAGAGCTGAGCCATGCGATGCGATGGTGGCCGCAATGACCGTGTGAG 765
Db 794 CAGAAAGTAGAGCTGAGCCATGCGATGCGATGGTGGCCGCAATGACCGTGTGAG 853
Qy 766 GCAGTATACCCAGATGAGAGAGTGTGCCAGACCATTTAGAGACACACCGGAGACAG 825
Db 854 AGCATATCTCTACAGTGGAGAGCTGTGAGAGTGCCAAAGAGAAATATCCACAGAGTG 913
Qy 826 AAGCAACTGTTAAACAGAGAGTTCAGAGACCCCTGTGGCGGCTTTTGGAGAGAGCCAGAGC 885
Db 914 AAGGAGAGAGCTAGTGAAGATTGACACCTCTTACGCCATCTGTGATGAGAAAGAGAGC 973
Qy 886 GAACGCTTCAAGCACTGGCCCGGAGAGAGAGAGAGTGTGACAGCGCGTGGCGGCGCTC 945
Db 974 GAGTGTCTGACAGCGATCAGCAGAGAGAGAGAGAGTGTGAGGCTTCAATCAGAGCTGTG 1033
Qy 946 ATCCGCCATGACGAGAGACCATTTGAGAGGCTCTCTCAAGCTGTGAGAGTCCGCCATCCAG 1005
Db 1034 ATCCCTCAGATACAGGAGAGAGCTGTGAAAGTCCACCAAGCTTGTGAGAGACCGCATCCAG 1093
Qy 1006 TCCATGAGAGAGAGCCAGATGTGCTCTCTACCTCCAGACAGAGCAAGAGTGTGAGAGTCCGCCATCCAG 1065
Db 1094 TCCCTGATGAGAGCCGAGAGAGGCTTCTCTCTCAAGTGCAGAGCATCATCAGAGAGC 1153
Qy 1066 GTCCGGCAATGTGCAAGGTGAGCTGCGAGAGAGAGCCGAGAGAGAGTGTGAGAGAGC 1125
Db 1154 ATTTAGAGAGAGCTTCCAGAGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1213
Qy 1126 GAGCAATCTCTGTGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1185
Db 1214 GAGCAATCTCTGTGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1273
Qy 1186 CCGGCGCGCGCTGGGGATG 1204
Db 1274 ACAGGTAAAGATGTGATG 1292

RESULT 6
LOCUS HSA291713 2097 bp mRNA linear PRI 06-JUL-2001
DEFINITION Homo sapiens mRNA for RNF28 gene for ring finger protein 28.
ACCESSION AJ291713.1 GI:13171050
VERSION AJ291713.1
KEYWORDS ring finger protein 28; RNF28 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2097)
AUTHORS Center,T., Iano,J., Kimura,E., McElhinny,A.S., Pelin,K.,
Witt,C.C., Bang,M.L., Tromblitas,K., Granzier,H., Gregorio,C.C.,
Sorninichl,H. and Labelt,S.
Identification of muscle specific ring finger proteins as potential
regulators of the titin kinase domain
J. Mol. Biol. 306 (4), 717-726 (2001)
JOURNAL MEDLINE 2110140
PUBMED 11243782
REFERENCE 2 (bases 1 to 2097)
AUTHORS Center,T.
TITLE Direct Submission
JOURNAL Submitted (12-FEB-2001) Center T., Structure and Biocomputing,
EMBL-Heidelberg, Meyerhofstrasse 1, Heidelberg 69112, Germany
REMARK Revised by author 22-FEB-2001
FEATURES
source
1..2097
/organism="Homo sapiens"
/misc_feature /db_xref="taxon:9606"
436..438
/note="putative alternative translation initiation codon"
475..1497
gene

```

```

CDS
/gene="RNF28"
475..1497
/gene="RNF28"
/note="short isoform"
/codon_start=1
/product="ring finger protein 28"
/protein_id="CAC3173.1"
/db_xref="GI:13171051"
/translation="MENLEKOLIPICLEMETKPYVLLPCOHLRCRCANDIFQANP
YMTSRGSSVSMGSRFCPTREHYIMRHHVYGLORMLVLENIIDYIKOECSSRPLQ
KGSHPMCKEHEKHEKINITYCLCEVPTMGKACFEVAPLOSFOGCKEELANC
ISMVAGNDRVOTIITOLEDSRRYTKREMSHVKELSKRFTLVAIDKESLQRI
TORERKSTFEALIOQYOBQDSDSTKIVETRAISLDEPGATFLIKOLIKSTVEA
SKGCQLKTEDEGFENMDFITLDLHIALRALIDFIDEDEEELFEEDDEEESTEG
KEEGHQ"
BASE COUNT 565 a 460 c 604 g 468 t
ORIGIN
Query Match 33.8%; Score 483.6; DB 9; Length 2097;
Best Local Similarity 66.7%; Pred. No. 7,1e-84;
Matches 706; Conservative 0; Mismatches 349; Indels 3; Gaps 1;
Qy 226 CTGCTAAGGGATGGGCACACATGAGACACTTGGAGAGAGAGCTATTGCCCATCTGC 285
Db 454 CTGATCCAGATGGGAATCCCATGAGAACTTGGAGAGAGAGAGCTATGCTGCCCTATCTGC 513
Qy 286 CTGAGATGTTCTCCAGAGCCGCTGTGATCTTCCCTGCCAAGACACACCTGTGCGCAG 345
Db 514 CTGGAGATGTTTACCAAGCCAGGTGATCTTGCCTGCCAGCACACCTGTGCGGAG 573
Qy 346 TGTGCCAAGAGCTCTTCCAGAGCTCTTATCTCTGTGGCAATCCCGGGGCTCCACACAG 405
Db 574 TGTGCCAATGATCATCTTCCAGAGCTGCAAAATCCCTACTGAGACAGCCGGGAGCTCATGTG 633
Qy 406 GTGCTTCCAGAGAGAGCTTCCGATGCCCATCTTGTAGACAGAGAGTGTCTGTGACAG 465
Db 634 TCCATGTCTGGAGGCGCTTCCCGCTGCCACCTGCGCCAGAGAGTATCATGATGTGT 693
Qy 466 CATGTGCTATGGCGCGACAGGAGACCTGTACTGTGAGAGAGAGATCATTTGACATCTACAG 525
Db 694 CAGGAGATGTAAGGCGCTGACAGAGAGAGCTGTGTGAGAGAGATCATGACATCTACAAA 753
Qy 526 CAGGAGTCTTCCCGGCGACATGACAGCCAGAGCTGAACAGCATCTGATGTGAGAGAGC 585
Db 754 CAGGAGTCTTCCAGGTGGCGGCTGCAGAGAGAGCT---CACCCATGTGCAAGAGAGC 810
Qy 586 GAGGAGAGAGATCAACATCTACTGCTTGAAGTGTGAGAGTGTGCCACCTGCTCTCTGC 645
Db 811 GAAGATGAGAAATCAACATCTACTGCTTGAAGTGTGAGAGTGTGCCACCTGCTCTCTGC 870
Qy 646 AAGGTTTGGCGCCGACAGAGAGTGTGAGTGGCGGCTGTGCCCATTTTACAAAGC 705
Db 871 AAGGTTTGGGATTCACAGAGAGGCTGTGAGAGTGGCGGCTGTGCCCATTTTACAAAGC 930
Qy 706 CAGAGAGTGAAGTGCATGCGATGCGATGCTGTGAGAGAGAGAGAGAGAGAGAGAGC 765
Db 931 CAAAGAGTGAAGTGAATACATGATCTCATCTGCTGTGAGAGAGAGAGAGAGAGAGC 990
Qy 766 GCAGTATACCCAGATGAGAGAGAGTGTGCCAAGACATTGAGAGACACACCGGAGACAG 825
Db 991 ACCATATCTACTAGTGTGAGATGCCCGTGCAGTGCACCAAGAGAGAGAGAGAGAGTA 1050
Qy 826 AAGCAACTGTTAAACAGAGAGTTCAGAGACCCCTGTGGCGGCTTTTGGAGAGAGCCAGAGC 885
Db 1051 AAGGAAGAGAGTGAAGAGAGTGTGACAGCTTGTATGCTTATGCTTCAATCAGAGAGCCTC 1110
Qy 886 GAACGCTTCAAGCACTGGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 945
Db 1111 GAGTGTCTGACAGCGATCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 1170
Qy 946 ATCCGCCATGAGAGAGACCATTTGAGAGGCTCTCAAGCTGTGAGAGTCCGCATCCAG 1005
Db 1171 ATCCAGAGTACAGAGAGAGAGTGCAGCAAGTCCACAAAGCTGTGAGAACTGCCATCCAG 1230

```

```
OY 1006 TCCATGAGAGAGCCGACATGCTCTACTCCAGCAAGAGAGGCTACACAG 1065
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1231 TCCCTGGAGAGCGCTGGGAGAGCCACTTCTTCTGACTGCGAGCACTCATCAAGC 1290
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1066 GTCGGGCAATGTCGAGAGTGGAGCTGGCAGCGCCGAGCCAGCTATGAGCATG 1125
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1291 ATTGTGGAAGCTTCCAGAGGCTGGGAGGAGACAGAGCGCTTTGAGAACATG 1350
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1126 GAGCAATTCTGTGAGCGCTGAGCAGCTGGCCGAATGTTGCGAACATCGACTTCAG 1185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1351 GACTTCTTACTTGTGATTTAGACACATAGCAGAGCCCTGAGACCCATTTGACTTTGGG 1410
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1186 CCGGGCGCCGCTGGGAGTGAAGAGATGACGACATGGCTTTGGATGGGAGGGCAAT 1245
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1411 ACAGTGTGAGGAAGAGAAATTCATTGGAAGAAAGATCAGAGAGAGAGAGTCCACA 1470
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1246 GCGGGCTGAGAGAGAGCGGCTGGACGTGCCAGAGAG 1283
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1471 GAAGGAGAGGAAGAGACACAGTAAAGAGCGTGGATG 1508
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
HSA276484 1756 bp mRNA linear PRI 04-JAN-2002
LOCUS HSA276484
DEFINITION Human mRNA for muscle specific RING finger 2 (MURF2) protein (MURF2
gene)
ACCESSION AJ276484
VERSION AJ276484.1 GI:18073355
KEYWORDS MURF2 protein; RING finger protein.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1756)
AUTHORS Stanchi F.
TITLE Characterisation of MURF2, a new muscle-specific RING finger
protein of the RBCC family that associates with microtubules
unpublished
2 (bases 1 to 1756)
JOURNLT Stanchi F.
REFERENCE Direct Submission
AUTHORS Stanchi F.
TITLE Submitted (20-MAR-2000) Stanchi F., CRIBI Biotechnology Centre,
JOURNAL Universita di Padova, Via G. Colombo 3, Padova, 35121, ITALY
FEATURES
SOURCE
1. 1756
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="p35.1-35.2"
/clone="STRAIT03036"
/tissue_type="skeletal muscle"
/clone_id="pEAK8"
/dev_stage="adult"
1. 1756
/gene="MURF2"
1. 282
/gene="MURF2"
/number=1
124. 1185
/gene="MURF2"
/codon_start=1
/product="muscle specific RING finger 2 (MURF2)"
/protein_id="CAC81706.1"
/db_xref="GI:18073356"
/translation="MDYKSLIDGNDPMENLEKOLIPICILEMTKRPVILLPOCHNIC
RKCANDIFOANPYMTSRSSVSMGGRFCPTCRHEVIMDRHCVYGLQRLVLENTI
DIYKDCSSRPLOKSGHPMKCEHEDENINYCLCEVPTCSMCVFGIHRACEVAPLQ
SVFOGKTEILNMCISMLVAGNDRVQITITLEDSDRYTKRNSHQVKSLELSQKPTLVA
IIDEKSELITIOEOKELSTIEALIOOYOLDRDSTKLVETALIOSLDEPGGATFL
LAKOLIKSIVESKGGOLGTEGPFNMDFFLIDELHIALDALAIDPFGTEDEBEDEFI
EERDDDEESTBEKEBEGHO"
283. 455
exon
```

```
/gene="MURF2"
/number=2
456. 624
/gene="MURF2"
exon
/number=3
625. 720
/gene="MURF2"
/number=4
725. 954
/gene="MURF2"
exon
/number=5
955. 977
/gene="MURF2"
/number=6
978. 1102
/gene="MURF2"
exon
/number=7
1103. 1174
/gene="MURF2"
/number=8
1175. 1756
/gene="MURF2"
exon
/number=9
1720. 1725
/gene="MURF2"
polyA_site
polyA_signal
BASE COUNT 466 a 398 c 516 g 376 t
ORIGIN
Query Match 33.7%; Score 482; DB 9; Length 1756;
Best Local Similarity 66.6%; Pred. No. 1.5e-83;
Matches 705; Conservative 0; Mismatches 350; Indels 3; Gaps 1;
OY 226 CTGCTAGGAGGCGCACACAACATGAGCACTTGAGAGACAGCTCATTTGCCCATCTGC 285
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 142 CTGATCCAGAGATGGGAATCCATGGAGACTTGGAGAGCAGCTGATCTGCCATATCTGC 201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 286 CTGAGATGTTCTCCAAAGCCGCTGGATCTTGCCCTGCCAACACAACCTGTGCCGAG 345
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 202 CTGGAGATGTTTACCAAGCAGAGGATGTCATCTTGCCCTGCACGACAACCTGTGCCGAG 261
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 346 TGTGCCAACGAGCTCTTCCAGGCGCTTAATCCCTGTGTGGCAATCCCGGGGCTCCACAG 405
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 262 TGTGCCAATGACATCTTCCAGGCTGCAATCCCTACTGGACCAAGCGGGGAGCTCAGTG 321
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 406 GTCTCTCAGGAGGAGCTTCCGATGCCCATCTTGTAGGACAGAGTTGCTGGACAGG 465
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 322 TCCATGTCTGGAGGCCGTTTCCGCTGCCCATCTGCCCTCAGAGGTGATGATGATCGT 381
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 466 CATGTGTCATATGCGCTGCAGCGGAACCTGCTAGTGGAGAACATTTGACATCTACAG 525
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 CACGGAGTGTAGGCGCTGCAGAGGAACCTGCTGGTGGAGAACATCGACATCTACAAA 441
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 526 CAGGAGTCTCCCGGCGCACTGCACGCGCAAGGCTGAACAGACACTCANTGTGTGGAGACAC 585
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 CAGGAGTCTCCAGTGGCGCGCTGCAGAAAGGCAGT--CACCCATGTCCAAAGAGACAC 498
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 586 GAGGAGGAAGATCAACATCTACTGCTGAGCTGGAGTGGCGGCACTGCTCTCTGCG 645
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 GAGATGAGAAATCAACATCTACTGCTGAGCTGTGAGGTGGCGGCACTGCTCATGTGC 558
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 646 AAGGTTTGGGCGCCCAAGAGAGTGTAGGTGGCGGCTTGGCCCATTTTACAAAGC 705
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 559 AAGGTGTTGGGATCCCAAGGCGCTGAGAGTGGGCCCATTTGAGAGTGTCTTCCAGGGA 618
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 706 CAGAAAGTGAAGTGAAGCATGTGCGATGCGATGCTGGTGGCGGCAATGACCGTGTGAG 765
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 CAAAAGACTGAATGTATATCTATCTCATGCTGTGGTGGGGAATGACCGGTGTGAG 678
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 766 GCAGTATCAACAGATGAGAGAGTGTGCCAGACATTGAGAGCAACAGCGGAGACAG 825
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 ACCATCATCATCAGCTGGAGAGATTCCTCCGTGAGTACCAAGAGGAAGATGACACAGATA 738
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

QY 826 AAGCACTGTTAAACACAGAGTTCGAGACCCGTGCGCGGTTTGGAGAGCGCAAGGCC 885
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 AAGGAAGAGCTGACGACAGAGTTTGACAGCTTATGCCATCTCGATGAGAGAAAAGT 798
QY 886 GAACTGCTTCAAGCACTGGCCCGGAGCAGAGAGAGAGTTCAGAGCCGTGCGGCGCTC 945
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 799 GAGTTGCTGACGCGGATCACGACGAGCAGAGAGAAAAGCTTGAAGCTTCATCGAGCCCTC 858
QY 946 ATCCGCCATGACGAGACACTTGGAGGGCTCCTCAAGTGGTGGAGTCCGCATCCAG 1005
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 859 ATCCAGCAGTACACAGAGAGCTGAGCAAGTCCACAAAGCTGTGAAAGTGCATCCAG 918
QY 1006 TCCATGAGAGAGCCGACATGCTCTCTACCTCCACAGCAGCAAGAGAGCTGATCAACAG 1065
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 919 TCCCTGACGAGCCTGGGGAGACCACTCTCTTGTACTGCCAAGCACTCATCAAAAAGC 978
QY 1066 GTCCGGGCAATGTCAAGTGGAGCTGCGCAGAGAGCCGCGAGGCTATGAGAGCATG 1125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 979 ATTGTGAAAGCTTCCAAAGGCGTCCAGCTGGGGAAAGACAGAGCGCTTGAAGACATG 1038
QY 1126 GAGCAATTCCTGTGAGCGTGGAGCACGCGCGAATGTGGAGACCATCGACTCCAG 1185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1039 GACTTCTTTACTTTTGATTTAGACACATAGACAGCCCTGAGAGCCATTTGAGTGGG 1098
QY 1186 CCGGCGCGCGCTGGGGATGAAAGAGATGACGACATGCTTTGGATGGGAGAGAGGCAAT 1245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1099 ACAGATGAGAAAGAGAGAAATTTCATTGAAGAAAGATCAGAGAAAGAGAGAGTCCACA 1158
QY 1246 GCGGGGCTGGAGAGAGAGCGGCTGCACGTGCCAGAAAG 1283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1159 GAAGGGAAGAAAGAGACACACGATGAGAGCTGATG 1196

```

```

RESULT 8
LOCUS AX274927 1764 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 192 from Patent WO012777.
ACCESSION AX274927
VERSION AX274927.1 GI:16547559
KEYWORDS
SOURCE
ORGANISM Homo sapiens
            human.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
AUTHORS Hillman,J.L., Baughn,M.R., Yue,H., Lal,P., Lu,D.A., Patterson,C.,
        Azimzai,Y., Bandman,O., Tang,Y.T., Mathur,P., Shah,P., Au-Young,J.
TITLE Transcription factors
JOURNAL Patent: WO 012777-A 192 04-OCT-2001;
        Incyte Genomics, Inc. (US)
FEATURES
    source
        1..1764
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /note="Incyte ID No: 3575519CB1"
BASE COUNT 473 a 399 c 516 g 376 t
ORIGIN

```

```

Query Match 33.7%: Score 482; DB 6; Length 1764;
Best Local Similarity 66.6%: Pred. No. 1.5e-83;
Matches 705; Conservative 0; Mismatches 83; Indels 3; Gaps 1;
QY 226 CTGCTAGGGATGCGCAACATGACACTTGGAGAGAGCTCATTTGGCCCATCTGC 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 141 CTGATCCAGATGGAGATCCCATGGAGAACTTGGAGAGACGACTATCTGCTATCTGC 200
QY 286 CTGGAATGTTTCCAAAGCCCGGTGTAATTCTGCCCTGCCAACACAACTGTGCCGCAAG 345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 CTGGAATGTTTCCAAAGCCCGGTGTAATTCTGCCCTGCCAACACAACTGTGCCGCAAG 260
QY 346 TGTGCAAGACGCTTCCAGGCGCTTAATCCCTGTGTGCAATCCCGGGGCTCCACACAG 405

```

```

Db 261 TGTGCAATATACATCTTCCAGGCTGCAAAATCCCTACTGAGACAGCGGGGAGCTCAGTG 320
QY 406 GTGTCTTCAAGAGAGAGCTTTCGATGCCCATCTTGTAGACAGAGAGTGTCTGGACAG 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 321 TCCATGTCTGAGAGCGGTTTCCGCTGCCACACCTGCGCCACAGAGTATCATGTGATCTG 380
QY 466 CATGTGTCTATGGCTTGCAGCCGAGCACTGTAGTGGAGAAATCATTTGACATTCACAG 525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 381 CACGAGAGTGTACGCGCTGACAGAGAGACCTGTGTGGAGAACATCATGACATCTACAAA 440
QY 526 CAGGATTCCTCCCGGCACATCGACAGCGAGCTGAACAGCACTCATGTGAGAGACAGC 585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 441 CAGGATGTCTCCAGCTGGCGCTGACAGAGGCGAGT---CACCCATGTGCAAGGACAC 497
QY 586 GAGGACGAGAAAGATCAACATCTACTGCTGAGTGCAGAGTGGCCACCTGCTCTCTGTC 645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 GAAGATGAGAAATCAACATCTACTGTCTACGTTGAGGTGCCACCTGCTCCATGTGC 557
QY 646 AAGCTTTTGGCGCCCAAGAGAGTGTAGGTGGCCCTCTGCCACCATTTACAAAGCC 705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 AAGTGTGTGGGATCCCAAGAGGCTGAGAGTGGCCCATTTGACAGAGTGTCTCCAGGGA 617
QY 706 CAGAAAGTGTAGCTGAGCGATGCGATGCGATGCTGGTGGCGGCAATGACCGTGTGAG 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 618 CAAGAGCTGAACTGATGATTAATCTGATCTCCATGCTGTGGTGGGGAATGACCGTGTGAG 677
QY 766 GCAATGATCAACCCAGATGAGAGAGTGTGCCAGACCATTTGAGAGCAACAGCCGAGAGC 825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 678 ACCATCATCTACATGAGCTGTGGAGATTCCTCCGATGATGACCAAGAGAACAGTACCAGTA 737
QY 826 AAGCAACTTAAACCAAGAGTTCGAGACCTGTGCGCGGTTTGGAGAGAGCCAAAGGCC 885
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 738 AAGGAAGAGCTGAGCCAGAAAGTTTGACACGTTGTATCCATCTCGATGAGAAAGAAAAGT 797
QY 886 GAACTGCTTCAAGCACTGTGCGCCCGGAGCAGAGAGAGAGTTCAGCGCGGTGCGGGGCTC 945
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 798 GAGTGTGCTGACAGCGGATCAACGACGAGAGAGAGAGAGAGAAAGCTTACCTTCATGAGGCCCTC 857
QY 946 ATCCGCACTACGAGAGACCACTTTGAGAGGCTCTCAAGTGTGTGATCCGCCATCCAG 1005
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 858 ATCCAGAGTACCAAGAGAGCAGCTGACAAAGTCCACAAAGCTGTGGAAAGTCCATCCAG 917
QY 1006 TCCATGAGAGAGCCGCAAGATGGCTCTACTACCTCCAGAGGCCAAAGAGCTGATCAACAG 1065
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 918 TCCCTGAGAGAGCTTGGGGAGGACACCTTCTTGACTCCCAAGCACTCATCAAAAAGC 977
QY 1066 GTCCGGGCAATGTGCAAGGTGAGCTGCGAGAGAGAGCGCGGAGCCAGCTATGAGAGCATG 1125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 978 ATTGTGAAAGCTTCCAAAGGCTGCGCAGCTGGGGAAGACAGAGAGGCTTTGANAACATG 1037
QY 1126 GAGCAATTCCTGTGAGCGTGGAGACGCTGGCCGAATGTGGAGAACCATCGACTCCAG 1185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1038 GACTTCTTTACTTTTGATTTAAGACACATAGCAGAGCCCTGAGAGCAATTTGATTTGGG 1097
QY 1186 CCGGCGCGCGCTGGGATGAAAGAGATGACAGATGAGTGTGGATGGGAGGAGGGCAAT 1245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1098 ACAGATGAGAAAGAGAGAAATTCATTGAAGAAAGAGATAGAGAAAGAGAGATCCACA 1157
QY 1246 GCGGGGCTGGAGAGAGAGCGGCTGACGTGCCAGAAAG 1283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1158 GAAGGGAAGAAAGAGACACACGATGAGAGCTGATG 1195

```

```

RESULT 9
LOCUS AK056942 1764 bp mRNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens CDNA FLJ2380 fls, clone SKMUS1000064, moderately
ACCESSION AK056942
VERSION AK056942.1 GI:16552479
KEYWORDS Oligo clapping; fls (full insert sequence);
        Homo sapiens skeletal muscle CDNA to mRNA, clone_lib:SKMUS1
SOURCE

```

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
1  
Oshima, A., Takahashi-Fuji, A., Tanase, T., Inose, N., Takeuchi, K.,  
Arita, M., Mushashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R.,  
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,  
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,  
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,  
Watanabe, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B.,  
Suzuki, Y., Sugano, S., Nagahori, K., Masuno, Y., Nagai, K. and  
Isogai, T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 1764)  
Isogai, T., Otsuki, T. and Sugiyama, T.  
Direct Submission  
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,  
Genomics Laboratory, 1332-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB) (supported by Japan  
key Technology Center etc.) 5' & 3' end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: RAB and  
HRI.

FEATURES  
source  
1. 1764  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="SKM0100064"  
/tissue\_type="skeletal muscle"  
/clone\_lib="SKM01"  
/note="cloning vector: PME18SFL3"  
137..1198  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAB71318.1"  
/db\_xref="GI:16552480"  
/translation="MDYKSSLIODGNPMENLEKOLIPICILEMFTKPVVILPCQHNLC  
RKANDIFOPANPYMTSGSSVMSGFRPCPTCRHIVIMDRHGVYGLQRMILVENII  
DIYKORCSRPLOKSGHPMKCKHEDEKINICLCEVPTCMGCGACAPLQ  
SVYQKQTELNKISIMLVACNDPVQITITQLEDSRPYKEMSHOVKEELSKQPLIYA  
LIDKESSELQRTQDEKELSTFEALIQIOEDLDSTKYVFAISLDPGGATFL  
LTKOLIKSIVEASKGQGLKTRQGFENMDFLIDLAIDALRAIDFGTDEEEEFIL  
FEEDDEEEESTEGKEEGH"

BASE COUNT  
466 a 402 c 517 g 379 t

ORIGIN  
Query Match 33.7%; Score 482; DB 9; Length 1764;  
Best Local Similarity 66.6%; Pred. No. 1.5e-83;  
Matches 705; Conservative 0; Mismatches 350; Indels 3; Gaps 1;

QY 226 CTGCTAGGAGTGGCGCAACATGGAGCAACTGGAGAGCACTCATTTGCCCATCTGC 285  
DB 155 CTGATCCAGATGGGAATCCATGAGAACTGGAGAACAGCTATCTGCCCTATCTGC 214  
QY 286 CTGAGAGATGTTCTCCAAAGCCGCTGTGATCTTGCCCTGCCAACAACACTGTGCCGCAAG 345  
DB 215 CTGAGATGTTTACCAACCAAGTGTGATCTTGCCCTGCCAACAACACTGTGCCGCAAG 274  
QY 346 TGTGCCAAGACGCTTCTCAGAGCCCTCTAATCCTCTGTGCCAATCCCGGGGGCTCCACAAG 405  
DB 275 TGTCCCAATGACATCTTCCAGGCTGCAATCCCTACTGAGACCAAGCCGGGGAGCTCAGTG 334  
QY 406 GTGCTTCAGAGAGAGCTTCCGATGCCATCTGTAGGACGAGAGTGTGCTGAGCAG 465  
DB 335 TCCATGTGTGAGAGCGCTTCCGCTGCCACCACTGCCGCGCAGAGAGTATTCATGATCGT 394  
QY 466 CATGGTGTCTATGGCCCTCAGCGAAGCTGTAGTGGAGAACATCATTCATCTCAAG 525

DB 395 CACGAGATTGAGGCTTCAGAGGAACCTGCTGTGGAGAACATCATCTACAA 454  
QY 526 CAGAGTCTCCCGCCACTGCAGCCCAAGCTGAACAGCACCCTATCTGTGTAGGAGCAG 585  
DB 455 CAGAGTCTCCAGTCCGCGCTGCAGAGAGGAGCAG - -CACCCCTGTGCAAGGAGCAG 511  
QY 586 GAGGAGGAGATCAACATCTACTGCTGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 645  
DB 512 GAAGTGTGAAATTAACATCTACTGCTGCTCAAGTGTGAGAGTGGAGTGGAGTGGAGTGGAG 571  
QY 646 AAGTGTTCGCGCCGACAGAGAGCTGTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 705  
DB 572 AAGGTGTTCGATCCAGAGAGCTGTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 631  
QY 706 CAGAGAGTGCAGCTGAGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 765  
DB 632 CAAAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 691  
QY 766 GCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 825  
DB 692 ACCATCATCTACAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 751  
QY 826 AAGCAACTGTAAACCAAGAGTTCGAGACCTGTGCGGCTTTTGGAGAGCCGCAAGGC 885  
DB 752 AAGGAAGAGCTGCAGCAGAGTTCGAGACGTTGATGCCATCTCGATGAGAGAAAGT 811  
QY 886 GAAGCTCTCAACCACTGAGCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 945  
DB 812 GAGTGTCTCAGCGGATCACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 871  
QY 946 ATCCGCGATACGAG 1005  
DB 872 ATCCAGCAGTACAG 931  
QY 1006 TCCATGAG 1065  
DB 932 TCCCTGAG 991  
QY 1066 GTCGAG 1125  
DB 992 ATTTGTGAG 1051  
QY 1126 GAGCAATTCCTGTGAG 1185  
DB 1052 GACTTCTTACTTGTGATTTAGACACATAGACAGCCCTGAGAGAGAGAGAGAGAGAGAG 1111  
QY 1186 CCGGCGCCCTGCGGATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1245  
DB 1112 ACAGATGAT 1171  
QY 1246 GCGGCGCTGAG 1283  
DB 1172 GAT 1209

RESULT 10  
AF353673 1791 bp mRNA linear PRI 25-APR-2001  
LOCUS AF353673 Homo sapiens irf1 ring finger protein mRNA, complete cds.  
DEFINITION AF353673  
ACCESSION AF353673.1 GI:13785923  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
1 (bases 1 to 1791)  
Wistow, G.  
IRF: A Novel Ring Finger Protein From Iris  
Unpublished  
2 (bases 1 to 1791)  
Wistow, G.



TITLE Direct Submission  
JOURNAL Submitted (27-FEB-2001) MSF, NFI, 6/331, NIH, Bethesda, MD 20892,  
USA

FEATURES  
source location/Qualifiers  
1..1791

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="1p33-p31"  
/tissue\_type="Iris"  
120..1181  
/note="TRP"  
/codon\_start=1  
/product="iris ring finger protein"  
/protein\_id="AAK39519.1"  
/translation="MDYKSLTQDGNPMNLEKOLICPICLEMTKPVVILPCQHNLC  
RKCAIDFOANPYTSSRGSSVSMGGRFCRHEVIMDRHGVGLDRLNLENI  
DIYQECSSRLQKSHPMCKRHEDEKINITYLCEVPKCKVGFHKKAEVAPLO  
SVFQOKTELNNCISMLVAGNDRVQIITQLDSRWYENSHQVEELSHKEDTLVA  
ILDERKSELQRIODEOKLKFIALIQOYEOELDKSRKLVETATIOSIDEFGATFL  
LTAOKLSIYVASKGCOLKTEQGFENMDFTLDELHTADALRAIDFQDEDEBEFI  
EEPOEEESTEGKEEGHQ"

CDS  
BASE COUNT 505 a 398 c 512 g 376 t  
ORIGIN

Query Match 33.7%; Score 482; DB 9; Length 1791;  
Best Local Similarity 66.6%; Pred. No. 1.5e-83;

Matches 705; Conservative 0; Mismatches 350; Indels 3; Gaps 1;

226 CTGAGAGGAGTGGCCACACATGGACAACTTGGAGAGCAGCTCATTTGCCCATCTGC 285  
138 CTGATCCAGAGTGGAAATCCCATGGAGAACTTGAGAAACAGCTGATCTGCCTATCTGC 197  
286 CTGAGAGTGTCTCCAAAGCCCGTGGATCTTGCCTGCCCAACACCTGGGCGGAG 345  
198 CTGAGATGTTTACCAACAGCAGTGTCTTGCCTGCCAGACAACTGTGCCGGAAG 257  
346 TGTGCCAACAGCAGCTTCTCAGAGCTCTTAATCTGTGGCAATCCCGGGCTCCACAGC 405  
258 TGTCCCAATGACATCTTCCAGAGCTGCAATCCCTACTGACAGCCGCGGAGCTCAGTG 317  
406 GTGCTTCAGAGAGCAGTTCCTCGATGCCATCTGTAGAGCAGAGTTTCTCTGAGAG 465  
318 TCCATGTCTGAGAGCCGTTCCGCTGCCACCTCCGCGCAGAGAGTGTATGATCTGT 377  
466 CATGTGTCTATGAGCTTCCAGAGCAGTGTATGAGAGAACTATGATCTTCAAG 525  
378 CACGAGAGTGTACGGCTGCAAGAGAACTGCTGGTGGAGAACTATGACATCTAACA 437  
526 CAGAGTCTCCGCGGCACTGCAGCCCAAGCTGAACAGCACTTATGTGAGAGCAG 585  
438 CAGAGTGTCTCACTGCGCGCTGCAAGAGGCACT--CACCCCATGTGCAAGAGCAG 494  
586 GAGAGCAGAGAACTATCACTTACTGCTGAGCTGAGAGTGGCCCACTGCTCTGTGC 645  
495 GAGAGTGAAGAAATCACTTACTGCTGCTGAGAGTGGCCCACTGCTCTGCTGCTGC 554  
646 AAGTTTTCGCGCCCAAGAGAGTGTGAGTGGCCCTGCTGCCCACTTATTAACAGC 705  
555 AAGTGTGTTGGAGTCCCAAGAGGCTGCGAGGTGGCCCATGCAAGTGTCTTCAGGGA 614  
706 CAGAGAGTGTGAGCTGAGAGTGGCATGGCTGCTGTTGGGGGCAATGACCGTGTGAG 765  
615 CAAAGAGTGAAGTGAATCTGATCTTCCATGTGTGGTGGGGGAAGACCGTGTGAG 674  
766 GCAGTGTATCACCAAGATGAGAGTGTGCCAGACATTTGAGAGCAACAGCCGCGAGAG 825  
675 ACCATCATCTACTAGCTGAGAGTATCCCGTGAATGACCAAGAGAACTACACAGGTA 734  
826 AACCACTGTAAACAGAGAGTTCGAGACCTGTGCGGGGTTTTGGAGAGCGCAAGGCG 885  
735 AAGGAGAGCTGAGCCAGAGATTTGACACGTTTATGATCCATCTGAGATGAGAGAAAGT 794

QY 886 GAATGCTTTCAAGCACTGGCCCGGAGAGAGAGAGAGTGGAGCGGTGCGGGGCTTC 945  
DB 795 GAGTTGCTGCAGCGGATCACCGAGAGCAGGAGAAACCTTACCTTATTCAGAGGCCCTC 854  
QY 946 ATCCGCCAGTACGAGACACCTTGGAGGCTCTCTCAAACTGGTGAATGCCATCCAG 1005  
DB 855 ATCCAGCAGTACGAGAGAGCTGAGCAAGTCCACAAACCTGGTGAACCTGCATCCAG 914  
QY 1006 TCCATGAGAGAGCCCGAGATGAGCTCTTACTCCAGCAGCAAGAGAGGCTGATCAACAG 1065  
DB 915 TCCCTGAGAGAGCGCTGGGGGAGCCACCTCTCTTACTGACCAAGCAATCTCAAAAGC 974  
QY 1066 GTCCGGGCAATGTCCAGAGTGAAGTGGCAGAGAGCGCGGAGCCAGCTATGAGAGCATG 1125  
DB 975 ATTTGGAAGTCTCCAAAGGCTGCCAGCTGGGGAGAGACAGACAGAGGCTTGAAGACATG 1034  
QY 1126 GAGCAATCTCTGTGAGGCTGGAGCAGCTGGCGGAATTTGCGAACCATTCGACTTCCAG 1185  
DB 1035 GACTTCTTACTTGTGATTTAGATAGACATAGCAGACGCCCTGAGAGCCATTAATCTTGGG 1094  
QY 1186 CCGGCGCCGCTGGGAGATGAAGAGATGACAGCATGGCTTTGATGGAGAGAGGCAAT 1245  
DB 1095 ACAGATGAGAGAGAGAGAGAAATTCATGAAGAAAGATTCAGAGAGAGAGAGTCCACA 1154  
QY 1246 GCGGGGCTGGAGAGAGAGAGGAGGAGCTGACCTGCCAGAGG 1283  
DB 1155 GAAGGAGAGAGAGAGAGACACCATTAAGAGCTGTGATG 1192

## RESULT 11

AY059627 1861 bp mRNA linear ROD 13-DEC-2001  
LOCUS  
DEFINITION Rattus norvegicus muscle ring finger protein 1 mRNA, complete cds.  
ACCESSION AY059627  
VERSION AY059627.1 GI:16444655  
KEYWORDS  
SOURCE Rattus norvegicus.  
ORGANISM Rattus norvegicus.  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

## REFERENCE

1 (bases 1 to 1861)  
Bodine,S.C., Latre,E., Baumhueter,S., Lai,Y.K.-M., Nunez,L.,  
Clarke,B.A., Poueymirou,W.T., Panaro,F.J., Na,E., Dharmarajan,K.,  
Pan,Z.-Q., Valenzuela,D.M., Dechata,T.M., Stitt,T.N.,  
Yancopoulos,G.D. and Glass,D.J.  
Identification of ubiquitin ligases required for skeletal muscle  
atrophy  
Science 294 (5547), 1704-1708 (2001)  
JOURNAL  
MEDLINE 21578247  
PUBMED 11679633  
REFERENCE  
2 (bases 1 to 1861)  
Nunez,L. and Glass,D.J.  
TITLE  
Direct Submission  
Submitted (12-OCT-2001) Muscle Research, Regeneron Pharmaceuticals,  
777 Old Saw Mill River Road, Tarrytown, NY 10591, USA

## FEATURES

source

## CDS

1..1861  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
95..1150  
/note="MurFL; monomeric E3 ubiquitin ligase"  
/codon\_start=1  
/product="muscle ring finger protein 1"  
/protein\_id="AA116405.1"  
/db\_xref="GI:16444656"  
/translation="MDYKSLIPDGNAMNLEKOLICPICLEMTKPVVILPCQHNLC  
RKCAIDFOANPYTSSRGSSVSMGGRFCRHEVIMDRHGVGLDRLNLENI  
DIYQECSSRLQKSHPMCKRHEDEKINITYLCEVPKCKVGFHKKAEVAPLO  
SVFQOKTELNNCISMLVAGNDRVQIITQLDSRWYENSHQVEELSHKEDTLVA  
ILDERKSELQRIODEKIDFIALILOYREOKSRKLVETATIOSIDEFGATFL







```

QY 667 GAGCTGTGAGGTGGCCCTCTGCGCCACCATTTACAAAGCCAGAGAGATGATGAGCAT 726
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 542 GAGCTGCCAGGTGGCTCCCTGACTCATGTCTTCCAGAGCAGAGATGATGATGAT 601
QY 727 GGCATCCGATGCTGTGTGGCGGCAATGACCGTGTGCGAGGATGATCACCAGATGAG 786
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 602 GGTATGCTGTACTGTGTGGGAGCAACGATAGATGATGATGATGATGATGATGAT 661
QY 787 GAGGTGTGCGCAGACCATTTAGAGCAACAGCCGAGCAGACGATGTTAAACAGAG 846
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 662 GACACCTGTAAACATTTATGATGCTGACAGAAACAGAAACAGACCTGTGTGAGAAA 721
QY 847 TTGCGAGCCCTGTGGCGGCTTTTGGAGAGAGCGGCAAGGCACTGTTCAAGCACTGGCC 906
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 722 TTTGATCACCTTATAGGCTCTCTGGAGAGAGAGAAACATGAAATGACCCAGCATCTACT 781
QY 907 CGGAGCAGAGAGAGAGAAATTTGACAGCGGCTGCGGGGCTCATCCGCCAGTACGAGACAC 966
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 782 CGAACACAGAGAGAGAAATGCAACATGTCGAACTTTATCAGAAAGTATTCGATCAC 841
QY 967 TTGAGAGGCTCTCTCAAGCTGTGTGAGTCCGCTCATGCTCATGAGAGAGCGGAGATG 1026
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 842 CTGGAGAAAGTATCCAAAGTGTGTGAGTCAAGATCCAGTTGATGATGAGCGCGAATG 901
QY 1027 GCTCTCTACTCTCCAGCAGCAGCAAGAGCTGATCAACAGAGTCCGGGCAATGTCGAGGTG 1086
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 902 GAGATTTTCTGACAGATGCCAAGACCTCTTGCAAAAGATGTCGAGAGCATCAAGCGG 961
QY 1087 GAGCTGGCAGAGAGCGCGGAGCCAGCTATGAGAGCATGAGCAATTTCTGTAGAGGTG 1146
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 962 TTTCAGATGAGAAATAGTAAACAGGTTATGATGATGATGATGATGATGATGATGATG 1021
QY 1147 GAGCAGCTGGCCAAATGTTGGCAACATGCTCCAGCGCGCGCGCGGAGTAA 1206
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1022 AATAGAGAAAGAAATTTATCCCTGAATTTGACTTTCTGAGAGAGAGAGAGAGAGAA 1081
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1207 GAGATGAGACATGCTTTGATGAGAGAGAGGCAATGCGGGGCTGTGAGAGAGAGCGG 1266
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1082 GATGCAAGAGAAATAGTGAAGAGAGAGAGAGAGAGATGATGATGATGATGATGATG 1141
QY 1267 CTGGAGCTGCCAGAGAGCTC 1286
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1142 GAAAAATGTTCAATATGATC 1161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
BC007750 1750 bp mRNA linear PRI 12-JUL-2001
LOCUS BC007750 Homo sapiens, clone MGC:12836 IMAGE:4110783, mRNA, complete cds.
DEFINITION BC007750 Homo sapiens, clone MGC:12836 IMAGE:4110783, mRNA, complete cds.
ACCESSION BC007750.1 GI:14043531
VERSION BC007750.1
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 1750)
Strausberg, R.
Direct Submission
Submitted (11-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcaps-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland:
Web site: http://www.nisc.nih.gov/

```

```

Contact: nisc_mgc@nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.T., Karlins, E., Legaspi, R.,
Lim, M., Maduro, O.L., Masilelo, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantirip, S., Thomas, P.J.,
Tlionson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: http://image.lnl.gov
Series: IRAL Plate: 17 Row: 4 Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.

FEATURES
    source
        1..1750
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="MGC:12836 IMAGE:4110783"
            /tissue="Muscle, Rhabdomyosarcoma"
            /clone_id="NHL_MGC_17"
            /lab_host="DH10B-R"
            /note="vector: pOTB7"
            190..1500
                /codon_start=1
                /product="Unknown (protein for MGC:12836)"
                /protein_id="AAH07750.1"
                /db_xref="GI:14043532"
                /translation="MDNLEKOLICPIEMFTRPVYLPCOHNLCRKASDIFQASNP
                TLPRGTTMASGGRFRRPSCHEEVYLDHRGXYGLORNLINIDIVROESTRPEKK
                SDPQCEHEERINITYCLNCEVPTCSLDCVGAHDCOVAPLTHFQKSLSTSGI
                AILVGSNDVQVLSQLEDTKTEIECKRQKQELCEKDYLYGLIEERKNEVYIT
                RTEEKLHEVRALIKRYSDBLHENVKLYSGIOFMDPEMAVFLQAKLLKLSIAS
                KAFOMEKIEHGVENHFNVLNREKIIREDIFYEDDEDEEBEGEBKEGEGV
                GEAVREVENYQTEFFGPDENPERKASELSOVELQAPALPVSSEPPALPPADA
                PVTQIGFEPPLQGGAAAPASSGSDSEARHIFSRWLSNLSNE"

BASE COUNT      556 a      389 c      433 g      372 t
ORIGIN
Query Match      31.5%; Score 450.6; DB 9; Length 1750;
Best Local Similarity 64.8%; Pred. No. 1.8e-77;
Matches 703; Conservative 0; Mismatches 374; Indels 8; Gaps 2;

QY 187 GAGCGCGCGGGGATGATGATTTACGGTGGTTCAGCGCTGCTAGGGGATGCCCAAC 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 GACAGCGAGAGATGAGAGCGCATCTGATTAACAAATCTTTCCAAAGAGACAGACAGCC 189
QY 247 ATGAGCAACTTGGAGAGAGAGAGATTTGCCCATGCGCTGGAGATGTTCCAGGCC 306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 190 ATGAGTAACTTAGAGAGAGAGATCTGTCCATCTGTGCTTAGAGATGTTACGAACCT 249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 GTGTGATCTTGGCCCTGCCAACACACCTGTGCCGAGAGTGTGCCAACAGAGCTTCCAG 366
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 250 GTGTGATTTCTCCCTGTGACAGCAACCTGTGTGAGAAATGTGCCAGATATTTCCAG 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 367 GCCTGTAACTCTGTGCGCAATCCCGGGGCTCCCAACAGGCTGTTTCAAGAGAGAGCTTC 426
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 310 GCCTGTAACTCTGTGCGCAATCCCGGGGCTCCCAACAGGCTGTTTCAAGAGAGAGCTTC 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 427 CGATGCCATCTTGTAGACAGAGAGTGTGCTGAGACAGAGATGTTGTTAGGCTGAG 486
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 370 CGCTGCCATCTGTGACAGAGAGAGTGTGCTGAGACAGAGATGTTGTTAGGCTGAG 429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 487 CGGAACTGTAGTGAAGAACATTTGATCTTACAGAGAGAGATCTCCCGGCACTG 546
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 430 AGGAACTGTGTTGGGAAATATATCATTTACATTAAGAGAGAGATCTCCCGGCACTG 486
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 547 CAGGCCAGGCTGAGAGAGAGATCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 606
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 487 ---GAAAAGAAATCGACAGAGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGATC 543
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 607 TACTGCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 544 TACTGCTGAACTGCGAAGTACCCACTGCTCTGTGCAAGGTGTTGGTGCACACAA 603
Oy 667 GACTGAGAGTGGCCCCCTTGGCCACCATTTCACAAAGCCAGAGAGTGGAGCGAT 726
Db 604 GACTGCGAGGTGGCTCCCTCTCACTGTTGTTCCAGAGAGAGTCTGAGCTCAGTAT 663
Oy 727 GGCATCGGATCTGTGGTGGCGGCAATGAGCCGTTGAGAGAGTGGATCACCAGATGGAG 786
Db 664 GGCATCGGATCTGTGGTGGCGGCAATGAGCCGTTGAGAGAGTGGATCACCAGATGGAG 723
Oy 787 GAGGTGCGCAACCACTTGTAGCAACAGCCGAGAGAGAGTGGATTTAAACAGAG 846
Db 724 GACACCTGCAAAACATATCATGCAATGTTGCAAGAAACAGAAAGCTTTGTGAGAG 783
Oy 847 TTGAGACCTCTGGCGGGTGGTGGAGAGCCGAGGCGCACTGCTTCAGAGCACTGGCC 906
Db 784 TTTGATTTACCTGTATGAGCTTTTGGAGAGAGAGAGAAATGAAATGACCAAGTCATTACC 843
Oy 907 CGGAGCAGAGAGAGAGAGTGTGAGCGCGTGGCGGCTTCATCCGCAATGAGAGAGCAG 966
Db 844 CGAAGCCCAAGAGAGAGAACTGGAACATGTCGCTGTGATCAAAAGTATTTGTATCAT 903
Oy 967 TTGAGGCGCTCTCAAACTGCTGTGAGTCCGCTATCCATGAGAGAGAGCCGAGAGT 1026
Db 904 TTGGAGAACGTTCTAAAGTGTGTTGATGAGTCAAGAAATTCAGTTATGATGAGCCAGAAATG 963
Oy 1027 GCTCTCTACTCTCAAGCAGAGAGAGAGTGTATCAACAAAGTGGGCGCAATGTCAGAGT 1086
Db 964 GCAGTGTCTTTCAGAGAAATGCCAAACCTGCTTAAATAAATTCGAGAACATCAAAAGGCA 1023
Oy 1087 GAGCTGCGAGAGAGCGCGGAGAGCGCTATGAGAGCATGAGCAATTCCTGTGAGCGT 1146
Db 1024 TTTGAGATGAGAAATATGAAACATGCTATGAGAACATGAAACACACTTACAGTCAACCTC 1083
Oy 1147 GAGCAGCTGGCGGAAATGTTGGAGCAATGCACTCCAGCCCGGCGGCTGGGGATGAA 1206
Db 1084 AATGAGAGAGAAAGATATATACGTAAATTTGACTTTTACAGAGAAAGTAAAGTAAAGAA 1143
Oy 1207 GAGGATGACGATGCTTTGATGAGAGAGAGCAATGCGGGCTGTGAGAGAGAGCGG 1266
Db 1144 GAGAGAGAGAGCG--GAGAGAGAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 1201
Oy 1267 CTGGA 1271
Db 1202 TAGAA 1206

RESULT 15
AK091728 1925 bp mRNA linear PRI 15-JUL-2002
LOCUS AK091728
DEFINITION Homo sapiens cDNA FLJ34409 f1s, clone HEART2001931, moderately
similar to Mus musculus RING-finger protein MORF mRNA.
ACCESSION AK091728.1 GI:21750167
VERSION AK091728.1
SOURCE Homo sapiens heart cDNA to mRNA, clone_11b:HEART2
clone:HEART2001931.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Nishi, T., Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,
Sugiyama, T., Ito, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Kanda, K., Magatsu, M., Murakawa, K., Kanehori, K.,
Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawahara, B., Suzuki, Y.,
Sugano, S., Nishihara, K., Masuko, Y., Nagai, K. and Isogai, T.
MEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1925)
AUTHORS Isogai, T. and Yamamoto, J.

```

```

TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomicshri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT MEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology (RAB); cDNA library
construction; Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing; HRI and
RAB; annotation; HRI and RAB.
FEATURES
source
1..1925
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEART2001931"
/tissue_type="heart"
/clone_lib="HEART2"
/note="cloning vector: PME18SFL3"
BASE COUNT 579 a 441 c 498 g 407 t
ORIGIN
Query Match 31.5%; Score 450.6; DB: 9; Length 1925;
Best Local Similarity 65.4%; Pred. No. 1.8e-77;
Matches 675; Conservative 0; Mismatches 351; Indels 6; Gaps 1;
Oy 187 GAGCGCGCGGAGTAACTTCAGGTGGTTCAGCCGCTGATGAGGATGCGCAAC 246
Db 35 GAGCAGAGAGAGAGAGAGTCTGTAATTAACAATCTTTTCCAAAGAGAGAGAGAC 94
Oy 247 ATGAGCACTTGGAGAGAGAGAGTATTTGCCCCATCTGCTGAGATTTTCCAGCCC 306
Db 95 ATGATATCTTGAAGAGAGAGAGAGTATGCTCCATCTGCTTGAAGATTTCCAGAACT 154
Oy 307 GTGGATCTTGGCTTCCCAACAGACCTGTCGCGAGTGGCAAGTGGCAAGGCTTCCAG 366
Db 155 GTGGATCTTCCCTTGTGAGCAACACTGTGTGAAGATTTGCAAGATTTTCCAG 214
Oy 367 GCTCTATCTCTGTGAGCAATCCCGGCTCCCAACAGGCTGTTCCAGAGAGAGTTC 426
Db 215 GCTCTATCTCTGTGAGCAATCCCGGCTCCCAACAGGCTGTTCCAGAGAGAGTTC 274
Oy 427 CGATGCCATCTTGTAGGAGAGAGTGTCTGAGAGAGAGTGTGATGAGGCTGAG 486
Db 275 CGTGGCCATCTGTGAGAGAGAGTGTGATGAGAGAGTGTGATGAGGCTGAG 334
Oy 487 CGAAGCTGCTAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
Db 335 AGGAGCTGCTGAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 391
Oy 547 CAGGCAAGGCTTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606
Db 392 ---GAAAGAAATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 448
Oy 607 TACTGCTGAGTGCAGAGTGCACCTGCTCTGTCGAGAGGTTTGGGGCCCAAG 666
Db 449 TACTGCTGAGTGCAGAGTGCACCTGCTCTGTCGAGAGGTTTGGGGCCCAAG 508
Oy 667 GACTGAGAGTGGCCCCCTTGGCCACCATTTCACAAAGAGAGAGAGAGAGAGAGAGAG 726
Db 509 GACTGAGAGTGGCCCCCTTGGCCACCATTTCACAAAGAGAGAGAGAGAGAGAGAGAG 568
Oy 727 GGCATCGGATCTGTGGTGGCGGCAATGAGCCGTTGAGAGAGTGGATCACCAGATGGAG 786
Db 569 GGCATCGGATCTGTGGTGGCGGCAATGAGCCGTTGAGAGAGTGGATCACCAGATGGAG 628
Oy 787 GAGGTGCGCAACCACTTGTAGCAACAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 846
Db 629 GACACCTGCAAAACATATCATGCAATGTTGCAAGAAACAGAAAGCTTTGTGAGAG 888
Oy 847 TTGAGACCTCTGGCGGGTGGTGGAGAGCCGAGGCGCACTGCTTCAAGCAGTGGCC 906

```

Db 689 TTGATACCTGATGCGATTTTGGAGGAGGAGGAATGAATGAATGACCAAGTCATTACC 748  
 QY 907 CGGAGCAGAGAGAGAGAGTTCAGCGCGTGGCGGGCTCATCCGCCAGTACGAGACAC 966  
 Db 749 CGAACCCAGAGAGAGAGAGTTCAGCGCGTGGCGGGCTCATCCGCCAGTACGAGACAC 808  
 QY 967 TTGGAGGCGCTCTCAAGCTGTGTGAGTCCGACATCCATCCATGAGGAGCGCGAGAT 1026  
 Db 809 TTGGAGAGAGCTCTCAAGTGTGTGAGTCCGACATCCATCCATGAGGAGCGCGAGAT 868  
 QY 1027 GCTCTCTACCTCCACAGGAGGAGGAGTATCAACAGGAGTGGGAGCAATGTGAGAGTG 1086  
 Db 869 GCAGTCTTTCTGCAAGAGGAGGAGGAGTATCAACAGGAGTGGGAGCAATGTGAGAGTG 928  
 QY 1087 GAGCTGCAAGAGGAGGAGGAGGAGTATGAGAGATGAGAGATTTCTGTGAGGCTG 1146  
 Db 929 TTTTCAGATGAGAGAGGAGGAGGAGTATGAGAGATGAGAGATTTCTGTGAGGCTG 988  
 QY 1147 GAGCAGGCTGCGGAGGAGGAGGAGTATGAGAGATGAGAGATTTCTGTGAGGCTG 1206  
 Db 989 AATGAGAGAGAGAGGAGGAGGAGTATGAGAGATGAGAGATTTCTGTGAGGCTG 1048  
 QY 1207 GAGATGAGAGGAG 1218  
 Db 1049 GAGAGAGAGGAG 1060

RESULT 16  
 HSA291712 2202 bp mRNA linear PRI 06-JUL-2001  
 LOCUS HSA291712  
 DEFINITION Homo sapiens mRNA for RNF29 gene for ring finger protein 29.  
 ACCESSION AJ291712.1 GI:13160384  
 VERSION AJ291712.1  
 KEYWORDS ring finger protein 29; RNF29 gene.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 2202)  
 Centner,T., Yano,J., Kimura,E., McElhinny,A.S., Pelin,K., Witt,C.C., Bang,M.L., Trombitas,K., Granzier,H., Gregorio,C.C., Sorimachi,H. and Lablanc,S.  
 Identification of muscle specific ring finger proteins as potential regulators of the titin kinase domain  
 J. Mol. Biol. 306 (4), 717-726 (2001)  
 JOURNAL  
 MEDLINE 2140140  
 PUBMED 11243782  
 2 (bases 1 to 2202)  
 Centner,T.  
 Direct Submission  
 Submitted (12-FEB-2001) Centner T., Structure and Biocomputing,  
 JOURNAL EMBL-Heidelberg, Meyerhofstrasse 1, Heidelberg 69112, Germany  
 REMARK Revised by author 22-FEB-2001  
 FEATURES  
 location/Qualifiers  
 1..2202  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 133..135  
 /note="putative alternative translation initiation codon"  
 /evidence=not\_experimental  
 181..1779  
 /gene="RNF29"  
 181..1779  
 /note="RNF29"  
 /note="alternative"  
 /codon\_start=1  
 /product="ring finger protein 29"  
 /protein\_id="CAC32840.1"  
 /db\_xref="GI:13160384"  
 /translation="MDNLEKOLICPTCLEMTFKPVLLPCOHNLCKKASDIFQASNP  
 YLPRTGRTMASGGRFCPCSRHEVYLDHRYGGLORLVLENIIDIKOESTREPK  
 SDPMCEHEERINITYCLNCEVPTSCSKFGAHHKQCVAPLTHVFOROKSEISDGI  
 AIIIVSGNDRVQGVISQLEDITCKTIECCRKOKOELCEKFDYLLGLERKKNEMTOYIT  
 AIIIVSGNDRVQGVISQLEDITCKTIECCRKOKOELCEKFDYLLGLERKKNEMTOYIT

CDS  
 join(181..1365,1654..1779)  
 /gene="RNF29"  
 /note="short isoform"  
 /codon\_start=1  
 /product="ring finger protein 29"  
 /protein\_id="CAC32839.1"  
 /db\_xref="GI:13160385"  
 /translation="MDNLEKOLICPTCLEMTFKPVLLPCOHNLCKKASDIFQASNP  
 YLPRTGRTMASGGRFCPCSRHEVYLDHRYGGLORLVLENIIDIKOESTREPK  
 SDPMCEHEERINITYCLNCEVPTSCSKFGAHHKQCVAPLTHVFOROKSEISDGI  
 AIIIVSGNDRVQGVISQLEDITCKTIECCRKOKOELCEKFDYLLGLERKKNEMTOYIT  
 RTOBEKLEHVRALIKKYSDDLNNKSKLVESGIGQPMDEBMAVFLQNAKTLTKISEAS  
 KAFOMEKIEHGYENMNHFTVNLNREKILIREIDYREDEDEEGBEGEGBEVG  
 GEAVEVEVENVOETPEDEBNPEKASELSOVELQAGALVSSPEPPALPPADA  
 PVTOIGFEPAPLOGQAAAPASGSDSEPARHIFSFWLNSLNE"

BASE COUNT 660 a 489 c 560 g 493 t  
 ORIGIN  
 Query Match 31.5%; Score 450.6; DB 9; Length 2202;  
 Best Local Similarity 64.8%; Pred. No. 1.8e-77;  
 Matches 703; Conservative 0; Mismatches 374; Indels 8; Gaps 2;  
 QY 187 GAGCCCGCGGGATGACCTTCAGCGGTTCACCGCGTGTAGGAGGATGCGCACAC 246  
 Db 121 GACACCGAGAGATGAGCCATCTGGAATTAACAATCTTTCCAAAGACGACGACAC 180  
 QY 247 ATGACAACTTGAGAGAGAGAGTTCATTCGCCATCTGCTGAGATGTTCTCCAGCC 306  
 Db 181 ATGATTAATTTAGAGAGAGAGAGTTCATTCGCCATCTGCTGAGATGTTCTCCAGAC 240  
 QY 307 GTGTGATCTTGGCTCCGCAACAACTGTGCGGCAAGTGTCCCAAGCAAGCTTCCAG 366  
 Db 241 GTGTGATCTTGGCTCCGCAACAACTGTGCGGCAAGTGTCCCAAGCAAGCTTCCAG 300  
 QY 367 GCCTTAATCTCTGTGGCAATCCGCGGCTCCACAAAGGTGCTTCCAGAGAGAGCTTC 426  
 Db 301 GCCTTAATCTCTGTGGCAATCCGCGGCTCCACAAAGGTGCTTCCAGAGAGAGCTTC 360  
 QY 427 CGATGCCATCTTGAAGAGAGAGGTGCTGAGAGAGAGAGTGTATGAGCTTCAG 486  
 Db 361 CGATGCCATCTTGAAGAGAGAGGTGCTGAGAGAGAGAGTGTATGAGCTTCAG 420  
 QY 487 CGAAGCTGTAGTGAAGAGAGAGTGTATGAGAGAGAGAGTGTATGAGCTTCAG 546  
 Db 421 AGAAGCTGTAGTGAAGAGAGAGTGTATGAGAGAGAGAGTGTATGAGCTTCAG 477  
 QY 547 CAGCCAGAGGCTGAACAGAGAGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGATC 606  
 Db 478 ---GAAAAGAAATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 534  
 QY 607 TACTGCTGAGAGAGAGAGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666  
 Db 535 TACTGCTGAGAGAGAGAGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 594  
 QY 667 GACTGTGAGAGAGAGAGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 726  
 Db 595 GACTGTGAGAGAGAGAGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654  
 QY 727 GGCATGCGAGAGAGAGAGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 786  
 Db 655 GGCATGCGAGAGAGAGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 714  
 QY 787 GAGGTGTGAGAGAGAGAGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 846  
 Db 715 GACACCTGCAAAATATGAGAGAGAGAGTGTATGAGAGAGAGAGAGAGAGAGAG 774  
 QY 847 TTGAGAGAGAGAGAGTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 906

Db 775 TTTGATTACCTGTATGCGATTTTGGAGAGAGAAAGTAATGACCAAGTCATTACC 834  
 Oy 907 CGGAGCAGAGAGAGAAAGTTGACGCCGCTGCGGGGCTCTATCCGCCAGTACGAGACCAC 966  
 Db 835 CGAACCAGAGAGAGAAAGTGGAGCAATGTCCTGCTGATCAAAAAGTATTCTGATCAT 894  
 Oy 967 TTGGAGGGCTCTCAAAAGCTGTGTGAGTCGCCATCCAGTCACATGGAGAGCCGACATG 1026  
 Db 895 TTGGAGAACGTCCTCAAAAGTTGGTTGATGAGTAATTCAGTTTATGATGAGCCAGGAATG 954  
 Oy 1027 GCTCTCTACCTCCAGCAGCAAGAGCTGATCAACAGAGTCGGGCAATGTCGAAAGTG 1086  
 Db 955 GCAGTGTCTCTCAGATATCCAAAACCTGCTGTAATAAAAATCTCGGAAGCATCAAGGCA 1014  
 Oy 1087 GAGTCGGCAGAGCGCGGAGCAGGCTGATGAGATGAGATGAGCAATTTCTGTGAGCGTG 1146  
 Db 1015 TTTTCAGATGAGAAATATAGACATGCTATGAGAACATGAAACACCTTCACAGTCACCTC 1074  
 Oy 1147 GAGCAGTGGCCGAAATGTTGGCAACATGACATTCACGCCGCGCGCTGGGGATGAA 1206  
 Db 1075 AATAGAGAAAGAAAGTATATAGCTGTAATTTGACTTTTACAGAAATGAAATGAAAGAA 1134  
 Oy 1207 GAGGATGACGACATGCTTGTGATGGGAGAGAGGCAATGCGGGCTGAGAGAGAGCGG 1266  
 Db 1135 GAGAGAGAGAGC--GGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1192  
 Oy 1267 CTGGA 1271  
 Db 1193 TAGAA 1197

RESULT 17  
 AK091310  
 LOCUS 2634 bp mRNA linear PRI 15-JUL-2002  
 DEFINITION Homo sapiens CDNA FLJ13991 fls, clone DFNES2007332, moderately similar to mus musculus RING-finger protein MORN mRNA.  
 ACCESSION AK091310  
 VERSION AK091310.1 GI:21749650  
 KEYWORDS oligo capping; fls (full insert sequence).  
 SOURCE Homo sapiens normal dermal fibroblasts (Neonatal Skin) (NHDF2564)  
 ORGANISM Homo sapiens  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1  
 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Waga, S., M., Murakawa, K., Kanehori, K., Takahashi, F., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuh, Y., Nagai, K. and Isogai, T.  
 NEDO human CDNA sequencing project  
 unpublished  
 2 (bases 1 to 2634)  
 Isogai, T. and Yamamoto, J.  
 Direct Submission  
 Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team): 2-6-7 Kazusa-Kametari, Kisarazu, Chiba 292-0812, Japan  
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
 NEDO human CDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan: CDNA full insert sequencing: Research Association for Biotechnology (RAB): CDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.): 5'- & 3'- and one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation: clone selection for full insert sequencing: HRI and RAB.

FEATURES  
 source  
 1. 2634  
 Location/Qualifiers  
 /organism="Homo sapiens"

Query Match 31.5%; Score 450.6; DB 9; Length 2634;  
 Best Local Similarity 64.8%; Pred. No. 1.8e-77;  
 Matches 703; Conservative 0; Mismatches 374; Indels 8; Gaps 2;  
 /db\_xref="taxon:9606"  
 /clone="DFNES2007332"  
 /cell\_type="normal dermal fibroblasts (Neonatal Skin) (NHDF2564)"  
 /clone\_11b="DFNES2"  
 /note="cloning vector: PMEB8FL3-primary culture, normal dermal fibroblasts"  
 BASE COUNT 793 a 580 c 586 g 675 t  
 ORIGIN  
 Oy 187 GAGCGCCGCGGAGTGAACCTTCACGGTGGGTTTCAAGCCGCTCTAGAGGATGCGACAC 246  
 Db 365 GACAGCGAGAGATGAGGCGCATCTGTAATTCGAATCTTTTCCAAAGAGCAGACACC 424  
 Oy 247 ATGGAACACTTGGAGAACAGCTCATTTGCCCATCTGCTGAGAGATGTTCTCAAGCCC 306  
 Db 425 ATGGAATTAATTGAGAAACCACTATCTGTCCTCATCTGTTAGAGATGTTCAAGAACT 484  
 Oy 307 GTGGTATCTTGCCTGCGCAACACCTGTGCGGAGTGGCCAGTGGCCAGACGCTTCCAG 366  
 Db 485 GTGGTATCTTCCCTTGTGACACACCTGTGTAGGAATGTGCCAGTGAATTTTCCAG 544  
 Oy 367 GCCTCTAATCTCTGTGGCAATCCCGGGCTCCACAGCGGTCTTCAAGAGAGCTTTC 426  
 Db 545 GCTCTAACCCTGTATTTGCCCCCAAGAGAGATACACCATGCGATCAGGGGCGGATTC 604  
 Oy 427 CGATGCCATCTGTGAGCAGAGGTTTCTGAGCAGGATGTTGTCTATGGCTTCAG 486  
 Db 605 GCGTCCATCTCTGTGACATGAAGTGTGTTGTAACATGAGGATGATGAGACTTCAG 664  
 Oy 487 CGGAACCTGTAGTGAAGAACATGATGATCTTACAGCAGAGAGATCCCGGCGCAGT 546  
 Db 665 AGGAACCTGTGTGAGAAATATCATTTGACATCTACAGAGAGAGATCCACAGGCA--- 721  
 Oy 547 CAGGCCAAGCTGAGAACAGCACCTCATGTGTAGAGACAGAGAGAGAGATCAATC 606  
 Db 722 --GAAAGAAATCCGACACAGCCCATGTGCGAGAACATGGAAGAGAGAGCATCAATC 778  
 Oy 607 TACTGCTGAGCTGCGAGAGTGGCCACCTGCTCTGCAAGGTTTGGGGCCACAGAG 666  
 Db 779 TACTGTGTGAACGTGCGAAGTACCATCTGCTCTGTGCAAGGTTTGGTGCACACAA 838  
 Oy 667 GACTGTGAGGTGGCCCTCTGCCACCATTTTCAACAGCCAGAGAGTGAAGTACAGCAT 726  
 Db 839 GACTGTGAGGTGGCTCCCTCCATCATGTGTTCAGAGACAGAGTGTGAGCTCAGTGAT 898  
 Oy 727 GGCATGCGCATGTGTGGGGGGGCAATGACCGTGTGAGCAGTGAATCACCAGATGGAG 786  
 Db 899 GGCATGCGCATCTCTGTGGCGAGAACGATGATGATCCAGGAGATGACGACCTGGAA 958  
 Oy 787 GAGGTGTGCCAGACCATTTGAGACCAACAGCCGACAGAGCAAGCAAGCTTTAAACAGAG 846  
 Db 959 GACACCTGCAAACTATGAGAGATGTGAGAAAGAAAGAAAGAGCTTTGTGAGAGAG 1018  
 Oy 847 TTGGAACCTGTGGCGGTTTGGAGAGCGAAGGGCAACTGTTCAAGCAAGCTGGCC 906  
 Db 1019 TTTGATTACCTGTATGATTTTGGAGAGAGAGAAATGAAATGACCCAAAGTATTACC 1078  
 Oy 907 CGGAGCAGAGAGAGAAAGTTGACGGCGTGGGGGCTCTCCGCCAGTACGAGACAC 966  
 Db 1079 CGAACCAGAGAGAGAACTGGAACATGTCCGCTCTGATCAAAAAGTATTGATCAT 1138  
 Oy 967 TTGGAGGGCTCTCAAAAGCTGTGAGTCCGCCATCCATGATGAGAGAGCCGACAGATG 1026  
 Db 1139 TTGGAGAACGTCCTCAAAAGTTGGTTGATGAGTAATTCAGTTTATGAGAGCAAGAAATG 1198  
 Oy 1027 GCTCTCTACCTCCAGCAGCAAGAGCTGATCAACAGAGTGGGGCAATGTGAGAGTG 1086  
 Db 1199 GCAGTGTCTCAGATATCCAAAACCTGCTGTAATAAAAATCTCGAAGAGCATCAAGGCA 1258









TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Perera, A., Shim, C., Thomas, E. and Kucherlapati, R.  
High Throughput Mouse Sequencing  
2 (bases 1 to 209211)  
Han, J., Montgomery, K.T., Grills, G., Chiu, D., Decker, J., Fusina, M.,  
Goltz, J., Halder, A., Hall, L., Ioshikhes, I.P., Lee, E., Long, J.,  
Perera, A., Shim, C., Thomas, E. and Kucherlapati, R.  
Direct Submission

TITLE  
JOURNAL

Submitted (29-NOV-2000) Department of Molecular Genetics, Albert  
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,  
Bronx, NY 10461, USA  
On Dec 21, 2001 this sequence version replaced gi:15148085.  
-----Genome Center

COMMENT

Center: Harvard Partners Genome Center  
Center Code: HPGC  
Web site: <http://www.hpcg.org/Sequence/mouse.html>  
Contact: [hpcg@mendel.mgh.harvard.edu](mailto:hpcg@mendel.mgh.harvard.edu)  
-----Summary Statistics

Center project name: ABD  
Sequencing vector: pUC18: L08752  
Chemistry: Dye-terminator Big Dye, 100%  
\*Consensus quality: 204023 at least Q20  
\*Consensus quality: 203326 at least Q30  
\*Consensus quality: 201766 at least Q40  
Estimated insert size: agarose-FP - N/A  
\*\*Estimated insert size: 208811 - sum-of-coverage

Quality coverage: agarose-FP - N/A  
Quality coverage: 10.2 x in Q20 bases; sum-of-coverage estimation

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 21 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```

1      30155: contig of 30155 bp in length
*      30156      30175: gap of unknown length
*      30176      76428: contig of 46253 bp in length
*      76429      76448: gap of unknown length
*      76449      99269: contig of 22821 bp in length
*      99270      99289: gap of unknown length
*      99290      119029: contig of 19740 bp in length
*      119030      119049: gap of unknown length
*      119050      133795: contig of 14746 bp in length
*      133796      133815: gap of unknown length
*      133816      147516: contig of 13701 bp in length
*      147517      147536: gap of unknown length
*      147537      161508: contig of 13972 bp in length
*      161509      161528: gap of unknown length
*      161529      171555: contig of 10027 bp in length
*      171556      171575: gap of unknown length
*      171576      182235: contig of 10660 bp in length
*      182236      182255: gap of unknown length
*      182256      190163: contig of 7908 bp in length
*      190164      190183: gap of unknown length
*      190184      194456: contig of 4273 bp in length
*      194457      194476: gap of unknown length
*      194477      199061: contig of 4585 bp in length
*      199062      199081: gap of unknown length
*      199082      201315: contig of 2234 bp in length
*      201316      201335: gap of unknown length
*      201336      201781: contig of 446 bp in length
*      201782      201801: gap of unknown length
*      201802      203287: contig of 1486 bp in length
*      203288      203307: gap of unknown length
*      203308      204141: contig of 834 bp in length
*      204142      204161: gap of unknown length
*      204162      204246: contig of 85 bp in length
*      204247      204266: gap of unknown length
*      204267      205695: contig of 1429 bp in length
*      205696      205715: gap of unknown length

```

# FEATURES

source

```

*      205716      206923: contig of 1208 bp in length
*      206924      206943: gap of unknown length
*      206944      207841: contig of 898 bp in length
*      207842      207861: gap of unknown length
*      207862      209211: contig of 1350 bp in length.
                        Location/Qualifiers
                        1. 209211
                        /organism="Mus musculus"
                        /strain="C57BL/6J"
                        /db_xref="taxon:10090"
                        /chromosome="5"
                        /clone="RP23-154K6"
                        /sex="male"
                        1. 30155
                        /note="assembly_name:Contig39"
                        30176..76428
                        /note="assembly_name:Contig38"
                        76449..99269
                        /note="assembly_name:Contig37"
                        99290..119029
                        /note="assembly_name:Contig36"
                        119050..133795
                        /note="assembly_name:Contig35"
                        clone_end:r7
                        vector_side:right"
                        133816..147516
                        /note="assembly_name:Contig34"
                        147537..161508
                        /note="assembly_name:Contig33"
                        161529..171555
                        /note="assembly_name:Contig32"
                        clone_end:SP6
                        vector_side:right"
                        171576..182235
                        /note="assembly_name:Contig31"
                        182256..190163
                        /note="assembly_name:Contig30"
                        190184..194456
                        /note="assembly_name:Contig29"
                        194477..199061
                        /note="assembly_name:Contig28"
                        199082..201315
                        /note="assembly_name:Contig27"
                        201336..201781
                        /note="assembly_name:Contig26"
                        201802..203287
                        /note="assembly_name:Contig25"
                        203308..204141
                        /note="assembly_name:Contig24"
                        204162..204246
                        /note="assembly_name:Contig23"
                        204267..205695
                        /note="assembly_name:Contig22"
                        205716..206923
                        /note="assembly_name:Contig21"
                        206944..207841
                        /note="assembly_name:Contig20"
                        207862..209211
                        /note="assembly_name:Contig19"

```

BASE COUNT 55539 a 49732 c 50169 g 53288 t 483 others  
ORIGIN

Query Match 25.6%; Score 367; DB 2; Length 209211;  
Best Local Similarity 100.0%; Pred. No. 2.3e-61;  
Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy      1 AAGGAGTAGACACAGAGTCTCGGAATATGACAGGGGTGAGAGAGCTTTAGGGGAGG 60
      |||||||
Db 190769 AAGGAGTAGACACAGAGTCTCGGAATATGACAGGGGTGAGAGAGCTTTAGGGGAGG 190828
Oy      61 GACAGGACTCTCCAGAGGAGGAGCAATAGCCGGGATCCCAAGATCCAGCAGCTTAAC 120
      |||||||
Db 190829 GACAGGACTCTCCAGAGGAGGAGCAATAGCCGGGATCCCAAGATCCAGCAGCTTAAC 190888

```

```
OY 121 TGACCGAGGAAGGTTGCACAGGCGAGGAGAAAGCCACAGACAGGGCCACAGCGAGCGAG 180
|||
Db 190889 TGACCGAGGAAGGTTGCACAGGCGAGGAGAAAGCCACAGACAGGGCCACAGCGAGCGAG 190948
|||
OY 181 GCTCCAGAGCGCGCGGGATGAACTTCACGCTGGGTTTCAAGCCGCTGCTAGGGGATGCG 240
|||
Db 190949 GCTCCAGAGCGCGCGGGATGAACTTCACGCTGGGTTTCAAGCCGCTGCTAGGGGATGCG 191008
|||
OY 241 CACAACATGGACAACTTGGAGAGCACTCATTTGCCCACTGCTGAGATGTTCTCC 300
|||
Db 191009 CACAACATGGACAACTTGGAGAGCACTCATTTGCCCACTGCTGAGATGTTCTCC 191068
|||
OY 301 AAGCCGCTGTGATCTTGCCTGCCCAACACAACTGTGGCGCAAGTGTGCCAAGAGCTC 360
|||
Db 191069 AAGCCGCTGTGATCTTGCCTGCCCAACACAACTGTGGCGCAAGTGTGCCAAGAGCTC 191128
|||
OY 361 TTCCAGG 367
|||
Db 191129 TTCCAGG 191135
```

Search completed: December 3, 2002, 10:09:28  
Job time : 3356 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 02:19:40 ; Search time 265 Seconds  
(without alignments)  
12160.791 Million cell updates/sec

Title: US-09-908-988b-1

Perfect score: 1 aagagagtgtagacagagtg.....ataagagactcaagtgtccc 1431

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseg 1011002.\*

1:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1431	100.0	1431	24	ABA99061	Murine muscle ring
2	924.2	64.6	1913	22	AAH68563	Human protein HP03
3	924.2	64.6	1990	22	AAH78026	Nucleotide sequenc
4	861.2	60.2	1349	21	AAH72433	Human nucleic acid
5	733.4	51.3	2040	24	ABN85313	Human cytoskeleton
6	626.4	43.8	1500	22	AAZ7653	DNA encoding human
7	485.2	33.9	2110	22	AA542490	Human cDNA encodin
8	484.6	33.9	1597	24	ABA99063	Murine muscle ring
9	482	33.7	1764	22	ABA83058	Human transcriptio

10	480.8	33.6	1231	22	AAH90037	Human bone marrow
11	478.8	33.5	1781	22	AA525842	Human cDNA encodin
12	450.8	31.5	2590	24	ABA99062	Murine muscle ring
13	439.6	30.7	1762	22	AA525855	Human cDNA encodin
14	419	29.3	1183	22	AAH90117	Human bone marrow
15	390.6	27.3	1796	22	AAH89924	Human bone marrow
16	267.4	18.7	531	21	AA525855	Human OREF ORF840
17	267.4	18.7	531	21	AA525855	Human interleukin
18	229.2	16.0	587	22	AA525314	Human cDNA encodin
19	162	11.3	446	22	AA525303	Human cDNA encodin
20	161.6	11.3	650	21	AA525760	Human OREF ORF1315
21	114.8	8.0	391	21	AA525969	Human OREF ORF1524
22	72.4	5.1	494	22	AA525969	Murine 7-transmemb
23	70	4.9	3262	21	AA525969	Human regulation f
24	66.6	4.7	799	19	AA525831	Nucleotide sequenc
25	66.6	4.7	1926	21	AA525831	Epstein Barr virus
26	66.6	4.7	1926	22	AA525902	Epstein Barr virus
27	66.6	4.7	2580	21	AA525902	Anti-sense strand
28	66.6	4.7	2580	24	AA525902	Vector phutidic DN
29	66.6	4.7	5452	20	AA525902	Vector phutidic DN
30	66.6	4.7	8705	20	AA525902	Nucleotide sequenc
31	66.6	4.7	9600	19	AA525902	Plasmid pCISBON f
32	66.6	4.7	10380	20	AA525902	Plasmid pCISBON f
33	66.6	4.7	10596	14	AA525902	Plasmid pCISBON f
34	66.6	4.7	10596	17	AA525902	Nucleotide sequenc
35	66.6	4.7	10596	20	AA525902	Nucleotide sequenc
36	66.6	4.7	16080	21	AA525902	DNA clone pCEK C1.
37	65	4.5	65	24	AA525902	Mouse spliced tran
38	63.8	4.5	795	22	AA525902	FGA insert stabl
39	63.6	4.4	825	22	AA525902	Human cDNA clone (
40	63.6	4.4	1738	22	AA525902	Human cDNA sequenc
41	63	4.4	1925	20	AA525902	Epstein Barr Virus
42	58	4.1	2004	18	AA525902	Nephila clavipes s
43	56	3.9	1588	22	AA525902	Spider silk protei
44	54.2	3.8	3489	21	AA525902	Kaposi's sarcoma-a
45	54.2	3.8	3489	22	AA525902	Nucleotide sequenc

# ALIGNMENTS

RESULT 1						
ID ABA99061 standard; DNA; 1431 BP.						
XX	ABA99061;					
AC	15-JUL-2002	(first entry)				
DT						
XX						
DE	Murine muscle ring finger protein 1 (MURF-1) coding sequence.					
XX						
KW	Muscle ring finger; MURF-1; mouse; cardiant; microtubule;					
KW	Intermediate filament; striated muscle; cardiac hypertrophy;					
KW	heart disease; gene; ds.					
XX						
OS	Mus musculus.					
XX						
FT	Key	Location/Qualifiers				
FT	CDS	199..1299				
FT		/*tag= a				
FT		/product= "MURF-1"				
XX						
PN	WO200206318-A2.					
XX						
PD	24-JAN-2002.					
XX						
PF	18-JUL-2001; 2001MO-US22896.					
XX						
PR	18-JUL-2000; 2000US-219020P.					
XX						
PA	(TEXA ) UNTIV TEXAS SYSTEM.					
XX						
PI	Olson EN, Spencer JA;					

XX WPI: 2002-241506/29.  
 DR P-PSDB: ABB08275.  
 XX  
 PT Novel muscle ring finger protein useful for drug screening, and for  
 diagnosing and treating diseases, particularly cardiomyopathies  
 XX  
 PS Claim 4; Page 123-125; 134pp; English.  
 XX  
 The sequence encodes murine muscle ring finger protein 1 (MURF-1). The  
 invention relates to a purified muscle ring finger (MURF) protein,  
 CC selected from MURF-1, MURF-2 and MURF-3. The MURF proteins of the  
 CC invention are involved in microtubule and intermediate filament  
 CC stabilisation of striated muscle cells and have cardiant activity.  
 CC The MURF proteins are useful for screening a candidate substance  
 CC for MURF protein-binding activity, in a cell, cell-free system or  
 CC in vivo, and its effect on interaction of MURF with microtubules,  
 CC homodimerisation of MURF, MURF-1, MURF-2 or MURF-3 stabilisation  
 CC of microtubules, interaction of MURF with intermediate filaments,  
 CC e.g. desmin, vimentin and cytokeratin, and heterodimerisation of MURF.  
 CC The screened compounds are useful for treating and preventing  
 CC cardiac hypertrophy and heart diseases. MURF proteins are useful as  
 CC antigens to immunise animals for the production of antibodies.

Sequence 1431 BP; 338 A; 384 C; 473 G; 236 T; 0 other:

Query Match 100.0%; Score 1431; DB 24; Length 1431;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-310;  
 Matches 1431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AAGGAGTGTAGACAGAGTGTGGAATAGACAGGGGTGAGAGAGCTGTAGGGGAAG 60  
 1 AAGGAGTGTAGACAGAGTGTGGAATAGACAGGGGTGAGAGAGCTGTAGGGGAAG 60  
 61 GACAGAGCTTCCAAAGAGGAGCAATAGCGGGATCCCAAGATCCAGTCAAGCTTAAC 120  
 61 GACAGAGCTTCCAAAGAGGAGCAATAGCGGGATCCCAAGATCCAGTCAAGCTTAAC 120  
 121 TGACCGAGAGAGGTGACAGAGGAGGAGCAAGCAAGAGGGCCACAGCGAGGAG 180  
 121 TGACCGAGAGAGGTGACAGAGGAGGAGCAAGCAAGAGGGCCACAGCGAGGAG 180  
 181 GTCGAGAGAGCGCGGGAGTGAACCTTACGGTGGTTCAAGCGCGTCTAGGGATCG 240  
 181 GTCGAGAGAGCGCGGGAGTGAACCTTACGGTGGTTCAAGCGCGTCTAGGGATCG 240  
 181 GTCGAGAGAGCGCGGGAGTGAACCTTACGGTGGTTCAAGCGCGTCTAGGGATCG 240  
 241 CACACATGAGACACTTGGAGAGAGCACTATTGGCCCATCTGCGAGATGTTCC 300  
 241 CACACATGAGACACTTGGAGAGAGCACTATTGGCCCATCTGCGAGATGTTCC 300  
 301 AAGCCGCTGGTGTCTTGGCCCTGCAACACAACTGTGCGGCAAGTGTGCCAGACGTC 360  
 301 AAGCCGCTGGTGTCTTGGCCCTGCAACACAACTGTGCGGCAAGTGTGCCAGACGTC 360  
 301 AAGCCGCTGGTGTCTTGGCCCTGCAACACAACTGTGCGGCAAGTGTGCCAGACGTC 360  
 361 TTCCAGAGCTTATCTCTGTGCAATCCGGGGCTCCCAACGGGTCTTCAGAGAGA 420  
 361 TTCCAGAGCTTATCTCTGTGCAATCCGGGGCTCCCAACGGGTCTTCAGAGAGA 420  
 421 CGTTCCGATGCCCATCTTGTAGGACAGAGGTTGCTCTGAGAGGATGTTATGAC 480  
 421 CGTTCCGATGCCCATCTTGTAGGACAGAGGTTGCTCTGAGAGGATGTTATGAC 480  
 481 CTGACAGCGAAGCTGTAGTGGAGACATCATGATCTACAGAGAGAGAGTCTCCCG 540  
 481 CTGACAGCGAAGCTGTAGTGGAGACATCATGATCTACAGAGAGAGAGTCTCCCG 540  
 541 CCAGTGCAGCGCAAGGCTTAACAGCACTATGTGAGAGAGCAGAGAGAGAGATC 600  
 541 CCAGTGCAGCGCAAGGCTTAACAGCACTATGTGAGAGAGCAGAGAGAGAGATC 600  
 601 AACATCTACTGCTGAGCTGCGAGAGTGCACCTGCTCTCTGCAAGGTTTTCGGGCC 660  
 601 AACATCTACTGCTGAGCTGCGAGAGTGCACCTGCTCTCTGCAAGGTTTTCGGGCC 660

QY 661 CACAAGAGCTGTAGGTGGCCCTCTGCCCCACCACTTACAAAGCCGAGAGTAGCTG 720  
 DB 661 CACAAGAGCTGTAGGTGGCCCTCTGCCCCACCACTTACAAAGCCGAGAGTAGCTG 720  
 QY 721 AGCATGGCATCCGGATGCTGTGGTGGCGGCAATACCGTGTGACAGGAGTACCCAG 780  
 DB 721 AGCATGGCATCCGGATGCTGTGGTGGCGGCAATACCGTGTGACAGGAGTACCCAG 780  
 QY 781 ATGAGAGAGTGTGCTCCAGACCATTTAGAGACAAAGCCGACAGCAACTGTTAAC 840  
 DB 781 ATGAGAGAGTGTGCTCCAGACCATTTAGAGACAAAGCCGACAGCAACTGTTAAC 840  
 QY 841 CAGAGTTCCAGACCCCTGTGCGGCTTTTGGAGAGCGCAAGGGCGAAGTCTTCAACA 900  
 DB 841 CAGAGTTCCAGACCCCTGTGCGGCTTTTGGAGAGCGCAAGGGCGAAGTCTTCAACA 900  
 QY 901 CTGGCCCGGAGAGCAGAGAGAGTTCAGCGCGTGGGGCTCTATCCGCACTACGGA 960  
 DB 901 CTGGCCCGGAGAGCAGAGAGAGTTCAGCGCGTGGGGCTCTATCCGCACTACGGA 960  
 QY 961 GACCACTTGGAGGGCTCTCAAAAGCTGTGAGTCCCGCATCCAGTCCATGAGAGAGCG 1020  
 DB 961 GACCACTTGGAGGGCTCTCAAAAGCTGTGAGTCCCGCATCCAGTCCATGAGAGAGCG 1020  
 QY 1021 CAGATGGCTCTTACCTCCAGCAGGCAAGAGCTGATCAACAAGTGGGGCAATGCG 1080  
 DB 1021 CAGATGGCTCTTACCTCCAGCAGGCAAGAGCTGATCAACAAGTGGGGCAATGCG 1080  
 QY 1081 AAGTGGAGCTGAGAGAGCGCGGAGCCAGGCTATGAGAGCATGAGAGCAATCTCTGTG 1140  
 DB 1081 AAGTGGAGCTGAGAGAGCGCGGAGCCAGGCTATGAGAGCATGAGAGCAATCTCTGTG 1140  
 QY 1141 AGCTGGAGCAGCTGGCCCAATGTTGCAACCATGACTTCCAGCCGGCGCGCTGGG 1200  
 DB 1141 AGCTGGAGCAGCTGGCCCAATGTTGCAACCATGACTTCCAGCCGGCGCGCTGGG 1200  
 QY 1201 GATGAAGAGATGACAGACATGCTTGGATGGGAGAGGAGGCAATGGGGCTGAGAGAG 1260  
 DB 1201 GATGAAGAGATGACAGACATGCTTGGATGGGAGAGGAGGCAATGGGGCTGAGAGAG 1260  
 QY 1261 GAGCGGCTGAGCTGCGAGAGAGCTCAGGCTCAGCTGACCGACACTGTATCCAGAGCG 1320  
 DB 1261 GAGCGGCTGAGCTGCGAGAGAGCTCAGGCTCAGCTGACCGACACTGTATCCAGAGCG 1320  
 QY 1321 ACACCGAAGCGGAGGCCAAGGATGCTGAGATCTGCGAGAGACACCGGCCACAA 1380  
 DB 1321 ACACCGAAGCGGAGGCCAAGGATGCTGAGATCTGCGAGAGACACCGGCCACAA 1380  
 QY 1381 GCTCGGCTTCCCGCCCGGAGAGTTCATTAAGGACTCAAGTGTCC 1431  
 DB 1381 GCTCGGCTTCCCGCCCGGAGAGTTCATTAAGGACTCAAGTGTCC 1431

RESULT 2  
 AAH68563  
 ID AAH68563 standard; cDNA; 1913 BP.  
 AC AAH68563;  
 DT 13-SEP-2001 (first entry)  
 XX Human protein HP03115 coding sequence.  
 DE Human; gene therapy; tumour; ss.  
 KW Homo sapiens.  
 OS Homo sapiens.  
 XX Homo sapiens.  
 PN WO200142302-A1.  
 PD 14-JUN-2001.  
 PF 06-DEC-2000; 2000WO-JP08631.

XX 06-DEC-1999; 99JP-0346863.  
 PR 06-DEC-1999; 99JP-0346864.  
 PR 08-FEB-2000; 2000JP-0031062.  
 PR 10-FEB-2000; 2000JP-0034090.  
 PR 10-FEB-2000; 2000JP-0034091.  
 PR 14-FEB-2000; 2000JP-0035829.  
 PR 14-FEB-2000; 2000JP-0035899.  
 PR 14-MAR-2000; 2000JP-0071161.  
 PR 30-MAY-2000; 2000JP-0160851.  
 PR  
 PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
 PI Kato S, Eguchi C, Saeki M;  
 XX WPI: 2001-381646/40.  
 DR P-PSDB: AAG93378.  
 XX  
 PT Human protein originated from tumor cell line, applicable as drug,  
 PT reagent for studying intracellular protein networks and protein source  
 PT for drug screening, also encoded cDNA for gene diagnosis and gene  
 PT therapy -  
 XX  
 PS Claim 3; Pages 236-240; 471pp; Japanese.  
 CC The present sequence is a human protein coding sequence. The human  
 CC protein, preferably originated from tumour cell line, is applicable as a  
 CC drug, a reagent for studying intracellular protein networks and a protein  
 CC source for screening proteins for binding low molecular weight drugs. The  
 CC human protein coding sequence is useful for gene diagnosis and gene  
 CC therapy, expression vectors and transformant cells for detection of  
 CC ligands and receptors.  
 XX  
 SQ Sequence 1913 BP; 446 A; 537 C; 612 G; 318 T; 0 other;  
 Query Match 64.6%; Score 924.2; DB 22; Length 1913;  
 Best Local Similarity 84.9%; Pred. No. 7.3e-197;  
 Matches 1117; Conservative 0; Mismatches 163; Indels 35; Gaps 6;  
 118 AACTGACCGAGGAGGATGACAGAGCAG-GGGAGAGGCGCAAGAGGCGCACAGCGAG 176  
 DB ATCTAAGCCAGAGGAGGATCTACAGGAGTGAAGGCCAGAGGAGGCGCGAGCGCAG 285  
 177 GGAGGCTCCAGAGCGCGCGGATGAACCTTCAGCGTGGGTTTCAAGCGCGCTCTAGGGGA 236  
 DB GCAGGAC-----ACGAGGAGGATGAACCTTCAGCGTGGGTTTCAAGCGCGCTCTAGGGGA 340  
 237 TGGCGACAAACATGACAACTTGGAGAGAGAGCTTCATTTGCCCATCTGCTGGAGATGTT 296  
 DB TGCACACAGCATGACAACTTGGAGAGAGAGCTTCATCTGCCCATCTGCTGGAGATGTT 400  
 297 CTCACAGCGCGTGTGATCTTGGCCCTGCCAACAACACCTGTCGGCAAGTGGCAACGA 356  
 DB CTCACAAACCACTGTGTATCTCTGCCCTGCCAACAACACCTGTCGGCAAGTGGCAACGA 460  
 401 CGTCTTCAGAGCGCTTAATCTCTGTGGCAATCCGGGGCTCCACAACGCTGTCTTCAGG 416  
 DB CGTCTTCAGAGCGCTTCGAATCTCTATGTGCAATCCGGGGCTCCACAACGCTGTCTTCAGG 520  
 461 AGGACGTTTCCGATGCGCATCTGTAGCAGAGAGTTCCTCTGACAGACAGCGGTGCTA 476  
 DB AGGACGTTTCCGATGCGCATCTGTAGCAGAGAGTTCCTCTGACAGACAGCGGTGCTA 580  
 521 AGGCGGTTTCCGATGCGCATCTGTAGCAGAGAGTTCCTCTGACAGACAGCGGTGCTA 580  
 477 TGGCCTTCAGAGCGGAACTGTAGTGAAGAACATCATTTGACATTCACAGAGGAGTCTTC 536  
 DB CGGCTTCAGAGCGGAACTGTAGTGAAGAACATCATTTGACATTCACAGAGGAGTCTTC 640  
 581 CGGCTTCAGAGCGGAACTGTAGTGAAGAACATCATTTGACATTCACAGAGGAGTCTTC 640  
 537 CCGGCACTGACAGCGCAAGGCTGAACAGACACCTCATGTGTAGAGAGCAGCAGAGAGAGA 596  
 DB CAGGCGGCTGACACCTCAAGGCTGAACAGCAGACCTCATGTGTAGAGAGCAGTGAAGAGAGA 700  
 641 GATCAATCATCTACGCTGTGAGCTGAGAGTGCACACCTGCTCTCTGCAAGGTTTGGG 656  
 597 GATCAATCATCTACGCTGTGAGCTGAGAGTGCACACCTGCTCTCTGCAAGGTTTGGG 656

DB 701 GATCAATCATCTACGCTGTGAGCTGAGAGTGCACACCTGCTCTCTGCAAGGTTTGGG 760  
 QY CGCCACAGAGAGCTGTAGGTGAGGCCCTCTGCGCACCATTTACAAAGCGCAGAGAGTGA 716  
 DB 761 TGGCCACAAGAGAGCTGTAGGTGAGGCCCTCTGCGCACCATTTACAAAGCGCAGAGAGTGA 820  
 QY 717 GCTGAGCGATGCGCATCGGATGCTGTGGGCGGCAATGACCGTGTGACAGCAGTATCAC 776  
 DB 821 GCTGAGCGATGCGCATCGGATGCTGTGGGCGGCAATGACCGCTGTGACAGCAGTATCAC 880  
 QY 777 CCAGTGTGAGAGAGTGTGCGCAGACCATTTAGAGCAACACCGCCAGCAGAGCAATGCTT 836  
 DB 881 ACAGATGTGAGAGAGTGTGCGCAGACCATTTAGAGCAACACCGCCAGCAGAGCAATGCTT 940  
 QY 837 AAACAGAGAGTGTGAGAGCGCTGTGCGCGGTTTGGAGAGCGCCAGAGGCGCAATGCTTGA 896  
 DB 941 AAACAGAGAGTGTGAGAGCGCTGTGCGCGGTTTGGAGAGCGCCAGAGGCGCAATGCTTGA 1000  
 QY 897 AGCACTGCGCGCGGAGCAGAGAGAGAGAGTTCAGAGCGCGTGTGCGCGGCTCTATCCGCACTA 956  
 DB 1001 GCGCGTGGCGCGGAGCAG 1060  
 QY 957 CGGAGACCACTTGGAGAGGCTCTCTCAAGCTGGTGAAGTCCGCGCATCAGTCCATGGAGGA 1016  
 DB 1061 TGGCGACCACTTGGAGAGGCTCTCTCAAGCTGGTGAAGTCCGCGCATCAGTCCATGGAGGA 1120  
 QY 1017 GCGCGAGATGCGCTCTCTACCTCCAGCAGCAGCAAGAGAGTGTATCAGAGAGTTCGGGCAAT 1076  
 DB 1121 GTCACAAATGCGCGCTGTATCTCAGCAGGCGCAAGAGAGTGTATCAGAGAGTTCGGGCAAT 1180  
 QY 1077 GTGAGAGTGTGAGCTGTGCGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1136  
 DB 1181 GTGAGAGTGTGAGCTGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1240  
 QY 1137 TGTGAGCGTGTGAG 1196  
 DB 1241 GCTAAGAGTGTGAG 1300  
 QY 1197 TGGGATGTGAGAGAGATGACAGACATGAGCTTTGATGGGAGAGAGAGAGAGAGAGAG 1256  
 DB 1301 CGGGAG 1354  
 QY 1257 GGAGAGAGCGGCTGTGAG 1316  
 DB 1355 GGAGAGAGCGGCGGAG 1408  
 QY 1317 GCGCACACCGGAG 1376  
 DB 1409 GCGCGGCTAG 1454  
 QY 1377 CCAGAGCTGCGCTTCCGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1431  
 DB 1455 CCAGAGCTGCGCTTCCGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1506  
 RESULT 3  
 ID AAH78026  
 AAH78026 standard; cDNA; 1990 BP.  
 XX  
 AC AAH78026;  
 DT 26-NOV-2001 (first entry)  
 XX  
 DE Nucleotide sequence of human cardiomyopathy associated protein (CAP).  
 XX Human: cardiomyopathy associated protein; CAP; myocardial biopsy;  
 KW left ventricular assist device; LVAD; cardiomyopathy;  
 KW ventricular tachyarrhythmia; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 283..1359



KM neurological disorder; cell proliferative disorder; cancer; tumour; ss.  
 XX Homo sapiens.  
 OS  
 XX MO200044900-A2.  
 PN  
 XX 03-AUG-2000.  
 PD  
 XX 28-JAN-2000; 2000MO-US02237.  
 PF  
 XX 29-JAN-1999; 99US-0117904.  
 PR 29-JAN-1999; 99US-0117905.  
 XX  
 PA (INCY-) INCYTE PHARM INC.  
 XX  
 PI Tang YT, Lal P, Hillman JL, Yue H, Azimzai Y, Lu AMD, Baughn MR;  
 PI Tran B, Shih LL, Au-Young JL;  
 XX  
 DR WPI: 2000-499332/44.  
 DR P-PSDB: AAB21048.  
 XX  
 PT Novel nucleic acid binding proteins, used to identify agonists and  
 PT antagonists of them, for the treatment of reproductive, immunological,  
 PT neurological and cell proliferative disorders including cancer -  
 XX  
 PS Claim 4: Page 178-179; 180pp; English.

Sequences AAA72382-A72436 represent cDNAs encoding novel human nucleic  
 acid-binding proteins (NuABPs: AAB20997-B21051). These cDNAs were  
 CC produced by extension from an appropriate EST (expressed sequence  
 CC tag) using primers designed using the EST. The invention also relates  
 CC to expression constructs, host cells and transgenic organisms comprising  
 CC a human NuABP nucleic acid, recombinant production of the human NuABPs,  
 CC and antibodies against the human NuABPs, and also to methods of  
 CC screening modulators of human NuABP activity or expression. The human  
 CC NuABPs, and their agonists and antagonists are used to treat diseases  
 CC associated with overexpression or underexpression of functional NuABPs.  
 CC Human NuABP proteins and nucleotides, and NuABP agonists and antagonists  
 CC can be used to diagnose, treat and prevent reproductive, immunological,  
 CC neurological and cell proliferative disorders. Reproductive disorders  
 CC that may be treated using compositions of the invention include  
 CC infertility, endometriosis, disruptions of the menstrual cycle and  
 CC disruptions of spermatogenesis. Immunological disorders that may be  
 CC treated include AIDS, allergies, and autoimmune disorders such as  
 CC multiple sclerosis, rheumatoid arthritis, diabetes and systemic lupus  
 CC erythematosus. Neurological disorders that may be treated include  
 CC epilepsy, neurodegenerative conditions such as Alzheimer's disease and  
 CC Parkinson's disease, prion diseases such as Creutzfeldt-Jakob disease,  
 CC and mental disorders such as schizophrenia. Cell proliferative disorders  
 CC that may be treated include a wide variety of cancers, and also  
 CC arteriosclerosis, atherosclerosis, cirrhosis and psoriasis.

Sequence 1349 BP; 328 A; 372 C; 428 G; 221 T; 0 other;

Query Match 60.28; Score 861.2; DB 21; Length 1349;  
 Best Local Similarity 88.28; Pred. No. 7.9e-183;  
 Matches 961; Conservative 0; Mismatches 123; Indels 6; Gaps 2;

QY 118 AACTGACCGAGGAGGTGCACAGCAG-GGAGAGGCGCAACGACAGGCGCCACAGCGAG 176  
 DB 52 ATCTAGCGAGGAGGCTCTACAGCAGTGTGAGGCCAGAGGAGGCGCCAGCGCCAG 111  
 QY 177 GCAGGCTCCAGAGCGCGCGGATGAACCTTCAGCGTGGTTTCAAGCCGCTGCTAGGGGA 236  
 DB 112 GCACGAGC-----ACGAGGGGATGAACCTTCAGCGTGGTTTCAAGCGCGCTGAGGGGA 166  
 QY 237 TGGCGACACATGACACACTTGGAGAGACAGCTATTTGCCCATCTGCTGAGATGTT 296  
 DB 167 TGCACACAGCATGTGACACACTGAGAGACAGCTATCTGCCCATCTGCTGAGATGTT 226  
 QY 297 CTCACAGCCGCTGGTGTCTTGGCTGCGACACACACTGCGCGCAGAGTGGCCACACA 356  
 DB 227 CTCACACAGCAGTGTGTCTCTGCGCTGCGACACACACTGCGCGCAATGTGCCACGA 286

QY 357 CGTCTTCAGGCGCTCTAATCCCTGTGGCAATCCGGGGCTCCACACAGGTTCTTCAGG 416  
 DB 287 CGTCTTCAGGCGCTCTAATCCCTGTGGCAATCCGGGGCTCCACACAGGTTCTTCAGG 346  
 QY 417 AGGACGTTTCCGATGATCCCATCTTGTAGGACAGAGTGTCTCTGACAGCATGTGTCTA 476  
 DB 347 AGGCGTTTCCGATGATCCCATCTTGTAGGACAGAGTGTCTCTGACAGCATGTGTCTA 406  
 QY 477 TGGCCTTCAGGCGGACCTGCTAGTGGAGAACATATGACATCTTCAAGCAGAGAGTCTC 536  
 DB 407 CGGCTTCAGGCGGAAACGCTGTAGTGGAGAACATATGACATCTTCAAGCAGAGAGTCTC 466  
 QY 537 CGGCGCTTCAGGCGGAAACGCTGTAGTGGAGAACATATGACATCTTCAAGCAGAGAGTCTC 596  
 DB 467 CAAGCGCTTCAGGCGGAAACGCTGTAGTGGAGAACATATGACATCTTCAAGCAGAGAGTCTC 526  
 QY 597 GATCAACATCTACTGCTGTAGCTGAGGTCGAGGTCGACCGCTCTCTCTGCAAGGTTTTCGG 656  
 DB 527 GATCAACATCTACTGCTGTAGCTGAGGTCGAGGTCGACCGCTCTCTCTGCAAGGTTTTCGG 586  
 QY 657 CGGCCACAGAGACCTGTAGGTCGAGGTCGACCGCTCTCTCTGCAAGGTTTTCGG 716  
 DB 587 TGGCCACAGAGACCTGTAGGTCGAGGTCGACCGCTCTCTCTGCAAGGTTTTCGG 646  
 QY 717 GCTGAGCGATGCGATCGCGATGCTGTGGGCGGCAATGACCGTGTGCGAGAGTATCAC 776  
 DB 647 GCTGAGCGATGCGATCGCGATGCTGTGGGCGGCAATGACCGTGTGCGAGAGTATCAC 706  
 QY 777 CGAGATGAGAGAGTGTGCGAGACCATTTGAGACACAGCCGCGACAGCAACTGTT 836  
 DB 707 ACAGATGAGAGAGTGTGCGAGACCATTTGAGACACAGCCGCGACAGCAACTGTT 766  
 QY 837 AAACCGAGGTTTGTAGAGCCTGTGCGGCTTTTGGAGAGGCGCAAGGCGCAACTGTTCA 896  
 DB 767 AAACCGAGGTTTGTAGAGCCTGTGCGGCTTTTGGAGAGGCGCAAGGCGCAACTGTTCA 826  
 QY 897 AGCAGTGGCGCGGAGAGAGAGAGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 956  
 DB 827 GGGCGTGGCGGCGGAGAGAGAGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 886  
 QY 957 CGGAGACCATTTGAGAGGCTCTTAAGCTGTGTGAGTGTGCGGCGGCGGCGGCGGCGGCGG 1016  
 DB 887 TGGCGACCATTTGAGAGGCTCTTAAGCTGTGTGAGTGTGCGGCGGCGGCGGCGGCGGCGG 946  
 QY 1017 GCGGCGATGCGGCTCTTAAGCTGTGTGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1076  
 DB 947 GCGGCGATGCGGCTCTTAAGCTGTGTGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1006  
 QY 1077 GTGAGAGTGTGAGTGTGAGAGAGAGAGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1136  
 DB 1007 GTGAGAGTGTGAGTGTGAGAGAGAGAGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1066  
 QY 1137 TGTGAGCTGTGAGTGTGAGAGAGAGAGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1196  
 DB 1067 GTTGTGAGCTGTGAGTGTGAGAGAGAGAGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1126  
 QY 1197 TGGGAGTGA 1206  
 DB 1127 CGGGGAGGA 1136

RESULT 5  
 ABN85313  
 ID ABN85313 standard; cDNA; 2040 BP.  
 XX  
 AC ABN85313;  
 AC  
 XX  
 DT 30-SEP-2002 (first entry)  
 XX  
 DE Human cytoskeleton-associated protein, CSAP-4, coding sequence.  
 XX  
 KW Human; cytoskeleton-associated protein; CSAP; CSAP-4;

KW cell proliferative disorder; viral infection; neurological disorder;  
KW transgenic animal; antiatherosclerotic; antipsoriatic; antiinflammatory;  
KW virucide; anticonvulsant; vasotropic; cerebroprotective; nootropic;  
KW neuroprotective; cytosolic; gene therapy; gene; ss.  
OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 291..1271  
FT /tag= a  
FT /product= "CSAP-4"  
PN W0200253719-A2.  
XX  
XX 11-JUL-2002.  
XX  
XX 04-JAN-2002; 2002W0-US00178.  
XX  
XX 04-JAN-2001; 2001US-260085P.  
PR 13-FEB-2001; 2001US-268554P.  
PR 14-FEB-2001; 2001US-269111P.  
PR 23-FEB-2001; 2001US-271211P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Lu DM, Baughn MR, Yao MG, Ding L, Honchell CD, Yue H, Tang YT;  
PI Warren BA, Dugan BM, Xu Y, Malia NK, Griffin JA, Stewart EA;  
PI Gandhi AR, Khan FA, Thangavelu K, Ison CH, Azimzal Y, Hatalia AD;  
PI Gietzen KJ, Lal PG, Sanjanwala MM, Elliott VS;  
DR P-SDB; ABB83475.  
XX  
XX MPI: 2002-583611/62.  
XX  
XX P-SDB; ABB83475.  
PT Novel isolated human cytoskeleton-associated protein for diagnosing,  
PT treating or preventing atherosclerosis, psoriasis, leukemia, epilepsy,  
PT ischemic cerebrovascular disease, cerebral neoplasms and Alzheimer's  
PT disease -  
XX  
XX  
XX Claim 5; Page 154-155; 167pp; English.  
XX  
XX The present sequence is the coding sequence for a human  
XX cytoskeleton-associated protein (CSAP). The CSAP and its coding sequence  
XX are useful in the diagnosis, treatment and prevention of a cell  
XX proliferative disorder such as actinic keratosis, atherosclerosis,  
XX psoriasis, primary thrombocythemia, leukamia; a viral infection such as  
XX those caused by adenoviruses (acute respiratory disease, pneumonia),  
XX atenoviruses (lymphocytic choriomeningitis); and a neurological disorder  
XX such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral  
XX neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease or  
XX amyotrophic lateral sclerosis. The CSAP coding sequence is also useful  
XX for creating knock out or knock in humanised animals or transgenic  
XX animals to model human diseases.  
XX  
XX  
XX Sequence 2040 BP; 465 A; 584 C; 650 G; 341 T; 0 other;  
SQ  
Query Match 51.3%; Score 733.4; DB 24; Length 2040;  
Best Local Similarity 78.2%; Pred. No. 3e-154;  
Matches 1028; Conservative 0; Mismatches 156; Indels 131; Gaps 7;  
QY 118 AACTGACGAGGAAGGTGACAGAGCAG-GGAGAGAGGCCACGACAGGCGCAG 176  
DB 214 ATCTAAGCGAGGAGGGGCTACAGCAGAGTAGAGGCGAGGAGCGCCAGCGCAG 273  
QY 177 GCAGGCTCAAGGCGCGCGGATGACACTTCACGCTGGGTTTAAAGCGCGTGAAGGGA 236  
DB 274 GCACGACC-----ACCGAGGGGATTAATTAATTCACAGTGGGTTTAAAGCGCTGAGGGA 328  
QY 237 TGCCGACACATGACACTGTTGAGAGAGCAGCTCATTTGCCCATCTGCTGAGATGT 296  
DB 329 TGACACACGATGAGCACTGAGAGAGCAGCTCATCTGCCCATCTGCTGAGATGT 388  
QY 297 CTCCAGCGCGTGTGATCTTCCCTGCCACACAACTGTGCGGCAAGTGTGCGCAAGA 356  
||||| ||| ||||||||| ||||||||||| ||||||||||| ||||||||||| |||||||||||

DB 389 CTCCAACCAAGTGTGATCTTCCCTGCCCAACACAACTGTGCGGCAAGTGTGCCAAGA 448  
QY 357 CGCTTCCAGGCGCTCAATCTCTGTGGCAATCCCGGGGCTCCACAAAGGCTTCCAGG 416  
DB 449 CGCTTCCAGGCGCTCAATCTCTGTGGCAATCCCGGGGCTCCACAAAGTGTGCCAAGA 508  
QY 417 AGGAGCTTTCCATGCGCATCTTGTGTAGGAGGAGGTTGTCTGTGACAGGAGTGTGCTA 476  
DB 509 AGGCGGTTTCCGCTGCCCATCTGTGAGGAGGATGAGGTTGTCTGTGACAGGAGTGTGCTA 568  
QY 477 TGGCTGACAGCGGAACCTCTGTAGTGAAGAACATTCATTCATCTACAAAGCAGAGTCTC 536  
DB 569 CGGCTGACAGCAAACTCTGTAGTGAAGAACATTCATTCATCTACAAAGCAGAGTCTC 628  
QY 537 CCGGCGACGACGACGCAAGGCTGACAGCAGCTCATGTGTGAGGAGGAGGAGGAGGAGAA 596  
DB 629 CAGGCGGCTGACCTCCAGGCTGAGAGCAGCTCATGTGTGAGGAGGAGGAGGAGGAGAA 688  
QY 597 GATCAACATCTACTGCTGAGCTGCGAGGTGCCACCTGCTCTCTGCAAGGTTTTCGG 656  
DB 689 GATCAATATTTACTGCTGAGCTGTGAGGTGCCACCTGCTCTCTGCAAGGTTTTCGG 748  
QY 657 CGCCCAAGAGACTGTGAGGTGCGCCCTCTGCGCCACATTTACAAAGCCGAAGAGTGA 716  
DB 749 TGCCCAAGAGACTGTGAGGTGCGCCCTCTGCGCCACATTTACAAAGCCGAGA----- 801  
QY 717 GGTGAGCATGCGATGCGCATGCTGTGGGCGGGAATGACCGTGTGACGAGTGTATAC 776  
DB 802 ----- 801  
QY 777 CCAGATGAGAGAGTGTGCGCAGACCATTTGAGAGACACAGCGCGAGAGCAACTGT 836  
DB 802 -----AGGACAAATAGCCGAGAGCAGAGCAAGCTGT 832  
QY 837 AAACCAAGAGTGTGAGACCTGTGCGGCTTTTGTGAGAGGCGCAAGGCGCAACTGCTTCA 896  
DB 833 AAACCAAGAGTGTGAGACCTGTGCGGCTTTTGTGAGAGGCGCAAGGCTGTGCTTCA 892  
QY 897 AGCAGTGGCGCGGAGCAGAGGAGAGAAAGTGTGAGCGGCTGTGCGGCTGTGCGGCTCA 956  
DB 893 GCGCGTGGCGCGGAGCAGAGGAGAGAAAGTGTGAGCGGCTGTGCGGCTGTGCGGCTCA 952  
QY 957 CGGAGACCACTTGTGAGGCTCTCAAGCTGTGTGAGTGTGCGGCTGTGCGGCTGTGCGGCTCA 1016  
DB 953 TGGCGACCACTTGTGAGGCTCTCAAGCTGTGTGAGTGTGCGGCTGTGCGGCTGTGCGGCTCA 1012  
QY 1017 GCGGCAATGCTCTTACTTCCAGAGGCGCAAGAGACTGATCAAAAGTGTGCGGCTCA 1076  
DB 1013 GCGCAAAATGCTCTTACTTCCAGAGGCGCAAGAGACTGATCAAAAGTGTGCGGCTCA 1072  
QY 1077 GTGGAAGTGTGAGCTGTGAGAGGCGCGGAGCGCGCTATGAGAGCATGAGCAATTTCTC 1136  
DB 1073 GTGGAAGTGTGAGCTGTGAGAGGCGCGGAGCGCGCTATGAGAGCATGAGCAATTTCTC 1132  
QY 1137 TGTGAGCTGTGAGACAGCTGTGCGCAAAATGTTTGCAGCAATGCACTTCCAGCGCGCGCCG 1196  
DB 1133 CTTAAGGCTGTGAGACAGCTGTGCGCAAAATGTTTGCAGCAATGCACTTCCAGCGCGCGCTT 1192  
QY 1197 TGGGATGGAAGAGATGAGACATGCTTGTGATGGGAGGAGAGGCAATGTGGGGCTGGA 1256  
DB 1193 CGGGAGGAGAGAGGAGG-----TGGCCCAAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGG 1246  
QY 1257 GAGAGAGCGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAGG 1316  
DB 1247 GGAAGAGCGGCGGAGTGGGCT-----TTAAGGCTGTGAGAGGCTGTGAGAGGCTGTGAGAG 1300  
QY 1317 GCGACACCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1376  
DB 1301 GCGCGGCTGTGAGAGTGGG-----GAGGATCTGCGGAGAGAGGAGGAGGAGGAGGAGGAG 1346  
QY 1377 CCAAGCTGCGCTTCCGCGCGCGGAGAGGTTCTCAATTAAGAGCTCAAGTGTCC 1431  
DB 1347 CCCAAATCGG---CGCGCGCGCGGAGGAGTCTCAATTAAGAGCTCAAGTGTCC 1398  
||||| ||| ||||||||| ||||||||||| ||||||||||| ||||||||||| |||||||||||



RESULT	6
AAAF27653	
ID	AAF27653 standard; DNA; 1500 BP.
XX	
AC	AAF27653:
XX	
DT	30-MAR-2001 (first entry)
XX	
DE	DNA encoding human transcriptional regulator protein #22.
XX	
KW	Human; transcriptional regulator protein; TXRCS; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200078954-A2.
PD	
XX	28-DEC-2000.
XX	
PF	15-JUN-2000; 2000MO-US16766.
XX	
PR	18-JUN-1999; 99US-0140109.
XX	
PA	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Lai P, Yue H, Tang YT, Baughn MR, Azimzai Y, Tran B;
XX	
DR	WPI; 2001-041425/05.
XX	
PT	Isolated polypeptide with a human transcriptional regulator protein
PT	sequence is useful for the diagnosis, prevention and treatment of
PT	disorders associated with the immune, reproductive and cardiovascular
PT	systems -
PS	
XX	Claim 5; Page 135-136; 142pp; English.
CC	The present invention relates to human transcriptional regulator
CC	protein (TXRCS) sequences. The antagonist and an agonist of the proteins
CC	of the invention are used to treat disorders associated with decreased
CC	or increased expression or activity of TXRCS.
XX	
SQ	Sequence 1500 BP; 401 A; 378 C; 439 G; 282 T; 0 other;
<hr/>	
Query Match	43.8%; Score 626.4; DB 22; Length 1500;
Best Local Similarity	87.5%; Pred. No. 2,2e-130;
Matches 721; Conservative	0; Mismatches 96; Indels 7; Gaps
<hr/>	
QY	118 AACGTGACCGAGGAAGGGTGCACAGCAG-GGGAGAAGGCCAACGACAGGGCCACAGCAG 176
DB	214 ACTCTAAGCAGGAAGGGTTACAGCGACGTGATGAAAGGCCAGAGGCCCCAGCCAG 273
QY	177 GCAGGCTTCAGAGCGCGCGGAGTGAACTTCAOGGTGGGTTTTCAAGCCGCTGTAGGGGA 236
DB	274 GCAGCGAC-----ACCGAAGGGAGTAGTAACACTTCAAGTGGGTTCAACCCTGCTAGGGGA 328
QY	237 TCGGCACAACATGAGAACAACTGGAGAGAGCATTTCCCCATCTGCTGGAGATGTT 296
DB	329 TGCACACACATGAGAACAACTGGAGAGAGCTCATCTGCCCATTGCTGGAGATGTT 388
QY	297 CTCGAAGCCCGTGTGATCTTGCCCTGCACACACAACCTGTGCCGGAAGTGTGCCAACA 356
DB	389 CTCGAAGCCCGTGTGATCTTGCCCTGCACACACAACCTGTGCCGGAAGTGTGCCAACA 448
QY	357 CGTCTTCAGAGGCTCTAAATCTCTGTGGCAATCCGGGGCTCCACAACGCTGTCTCAGG 416
DB	449 CGTCTTCAGAGGCTCCAATCTCTAAAGGAGTCCCGGGGCTCACACACTGTGCTTCAGG 508
QY	417 AGGACCTTTCGATGGCCAACTGTAGAGCAGAGGTTGTCTTGGAAGGCATGTTCTTA 476
DB	509 AGGCGCTTTCGATGGCCAACTGTAGAGCAGAGGTTGTCTTGGAAGCAGAGGTTCTTA 568
QY	477 TGCGCTGCAGCGGAACCTGCTAGTGGAGAACAATCATTTGACATCTACAGCAGAGTCTTC 536

Db	569	CGGCTTGCAGGAAAACCTGCTAGTGTGGAGACATTATTCGACATTTTACAAGCGAGAGTATC	628
QY	537	CCGGCCACTGCAAGCCCAAGAGCTGAAACAGCACCTCATGTGTGAGGAGCAGAGAGAA	596
Db	629	CAGGCCGCTGCACCTCAAGAGCGCTGAGACAGCACCTCATGTGCGAGAGCATGAAGAGAA	688
QY	557	GATCAACATCTACTAGTCCCTGAGACTGCGAGAGTCCCACTGCTCTCTGCAAGTTTTCG	656
Db	689	GATCAATATTATTACGTCCCTGAGACTGTGAGAGTCCCACTGCTCTCTCTGCAAGTTCG	748
QY	657	CGCCACAAAGACACTGTGAGTGTGCGCCCTCTGCCACATTACAAAGCCGCAAGAGTGA	716
Db	749	TGCCACAAAGACTGTGAGTGTGCGCCCTCTGCCACATTACAAAGCCGCAAGAGTGA	808
QY	717	GCTGAGCATGGCATCGGATGCTGTGTGGCGGCAATGACCGTGTGCAAGGATGATCAC	776
Db	809	GCTGAGCATGGCATCGGATGCTGTGTGGCGGCAATGACCGCGTGAAGCATGATCAC	868
QY	777	CCAGATGAGAGAGTGTGTCCACAGACCATTTGAGAGACAAAGCCGCAAGCAACTGTT	836
Db	869	ACAGATGAGAGAGTGTGTCCACAGACTTATGAGAGACAAATGACCGAGGCAAGTGT	928
QY	837	AAACGAGAGTTCGAGACCCCTGTGGCGGTTTTTGAGAGAGCGGACGGAAGTCTCTCA	896
Db	929	AAACGAGAGTTCGAGACCTGTGTGGCGAGTGTGAGAGAGCGGACGGAAGTCTCTCA	988
QY	897	AGCACTGGCCCGGGAGC-AGGAGGAGAAATTGACAGCGGTGGCG	939
Db	989	GCGGCTGGCCCGGAAACAAGCAGGACGACAAAGCTTCACAGCGGATCCG	1032
RESULT 7			
AA542490			
ID	AA542490 standard; cDNA; 2110 BP.		
AC	AA542490;		
XX			
DT	18-DEC-2001 (first entry)		
XX			
DE	Human cDNA encoding an mdt protein, clone LG:247384.1:200MAY19.		
XX			
KW	Human; molecules for disease detection and treatment; mdt; ss;		
KW	Antiarteriosclerotic; hepatotropic; antipsoriatic; cytostatic;		
KW	Immunosuppressive; antidiabetic; antiallergic; neuroprotective;		
KW	osteopthic; antiarthritic; cell proliferative disorder;		
KW	arteriosclerosis; cirrhosis; psoriasis; cancer; adenocarcinoma;		
KW	leukaemia; breast cancer; autoimmune disorder; AIDS;		
KW	acquired immunodeficiency syndrome; Addison's disease;		
KW	diabetes mellitus; asthma; multiple sclerosis; osteoarthritis.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200162922-A2.		
XX			
PD	30-AUG-2001.		
XX			
PF	21-FEB-2001; 2001WO-US05896.		
XX			
PR	24-FEB-2000; 2000US-0185213.		
PR	16-MAY-2000; 2000US-0205232.		
PR	17-MAY-2000; 2000US-0205285.		
PR	17-MAY-2000; 2000US-0205286.		
PR	17-MAY-2000; 2000US-0205287.		
PR	17-MAY-2000; 2000US-0205323.		
PR	17-MAY-2000; 2000US-0205324.		
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
XX			
PI	Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;		
PI	Chen A, d'Sa SA, Amshy S, Dahl CR, Dam TC, Daniels SE;		
PI	Dufour GE, Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL;		
PI	Liu TF, Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Daffo A,		

PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;  
 PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;  
 DR MPI: 2001-570631/64.  
 P-PSDB: AMU25438.  
 XX  
 PT New disease detection and treatment molecule polynucleotides and  
 PT polypeptides, useful for diagnosis and treatment of arteriosclerosis,  
 PT cirrhosis, psoriasis, cancer, autoimmune disorders, diabetes mellitus,  
 PT asthma and multiple sclerosis  
 XX  
 PS Claim 1: Page 121: 183bp: English.  
 XX  
 CC The invention relates to novel human molecules for disease  
 CC detection and treatment (mddt proteins) and the polynucleotides encoding  
 CC them. The MDDT polynucleotides and polypeptides are useful for diagnostic  
 CC and therapeutic purposes e.g. to diagnose and treat cell proliferative  
 CC disorders (e.g. arteriosclerosis, cirrhosis and treat cell proliferative  
 CC adenocarcinoma, leukaemia and breast cancer) autoimmune disorders  
 CC (e.g. acquired immunodeficiency syndrome (AIDS) and Addison's disease)  
 CC diabetes mellitus, asthma, multiple sclerosis, osteoarthritis, and many  
 CC more diseases given in the specification. The present sequence  
 CC encodes an mddt protein of the invention.

Sequence 2110 BP: 562 A: 465 C: 613 G: 470 T: 0 other;

Query Match 33.9%; Score 485.2; DB 22; Length 2110;

Best Local Similarity 66.8%; Pred. No. 8.2e-99; Matches 707; Conservative 0; Mismatches 348; Indels 3; Gaps 1;

QY 226 CTGCTAGGGGAGTCCGACACATGACATGAGACAGCTCATTTGCCCATCTGC 285  
 DB 143 CTGATCCAGGATGGGAATCCCATGGAGACTTGGAGAGAGCTGATCTCCATCTGC 202  
 QY 286 CTGAGATGTTCTCCAAAGCCCGGTGATCTTGCCCTGCCAACAACCGTGGCCGCAAG 345  
 DB 203 CTGAGATGTTTACCAACGCACTGTCTCTGCCAGACACACTGTGGCGGAG 262  
 QY 346 TGTGCCAAGCAGCTCTCCAGGCTCTAATCTGTGTGCAATCCCGGGGCTCCACAAG 405  
 DB 263 TGTGCCAATGATATTTTCCAGGCTCTAACCCTGATTTTGGCCCAAGAGAGTACCACC 322  
 QY 406 GTGTCTCAGAGAGACGTTCCGATGCCCATCTTGTAGGACAGAGTGTGCTTGACAG 465  
 DB 323 ATGCAATCAGGGGCGGCTCCGCTGCTCCATCTGTAGACATGAATGCTTTGGATAGA 382  
 QY 466 CATGCTCTAATGCTCCAGCTGCAAGCTGTAGTGAAGACATCTTGCATCTACAG 525  
 DB 383 CATGGGATATGACCTTCAAGAGAACTGCTGTGTGAGAAACATCTCAGATCTACAAA 442  
 QY 526 CAGGATGCTCTCCGCGCAGCTGACGCAAGGCTGAACAGCACCTCATGTGTAGAGAC 585  
 DB 443 CAGGATGCTCTCCAGCTGCGCCCTGCGAGAAAGGCACT--CACCCCATGTGCAAGAGCAC 499  
 QY 586 GAGACGAGAAAGATCAATCATCTAGCTGAGCTGAGAGTGGAGCCCATCTGCTCTGC 645  
 DB 500 GAGATAGAGAAATCAACATCTACTGTCTACGTGTGAGTGCCTCCACCTGCTCCAGTGC 559  
 QY 646 AAGTTTTCGCGCCCAACAGAGCTGTGAGTGGCCCTCTCCCACTTTCACAAAGC 705  
 DB 560 AAGTGTGTTGGATCCACAAGGCTGCGAGGTGGCCCATTCAGAGTGTCTTCCAGGGA 619  
 QY 706 CAGAAAGTGAAGCTGAGCGATGCGATGCGATGTGTGTGGCGGCAATACCGTGTGAC 765  
 DB 620 CAAAGACTGAACTGAATTAATCTGATCTCCATGTGTGTGGCGGGAATACCGTGTGAC 679  
 QY 766 GCAGTATCAACCCAGATGAGAGAGGTGTGCAGACATTTGAGAGCAACAGCCGCAACAG 825  
 DB 680 ACCATATCTACTACGCTGAGGATTCCTCGAGTGACCAAGAGAACGTACACAGGTA 739  
 QY 826 AAGCAACTGTTAAACAGAGAGTGTGAGACCTGTGCGGGTGTGGAGAGGACGAGGCG 885  
 DB 740 AAGGAGAGCTGAGCAGAGAGTGTGACACGTTGATGCACTCTGATGACAGAAAGAAAGT 799

QY 886 GAACGTCTTCAAGCACTGGCCCGGAGCAGAGAGAAAGTTGACCGCGTGGCGGCTTC 945  
 DB 800 GAGTTGCTGAGCGGATCACGACAGAGAGAGAAAGTTAGCTTCAATGAGCGCTTC 859  
 QY 946 ATCCGCGCAGAGAGACCACTTGGAGGCTTCCTCAAGCTGTGGAGTCCGCAATCCAG 1005  
 DB 860 ATCCAGAGATACCGAGAGCACTGTGACATTCACAAAGCTGTGTGAATCTCCATCCAG 919  
 QY 1006 TCCATGAGGAGCGCCGAGATGCTCTTACCTCCAGCAGGCAAGAGAGCTGTCAACAG 1065  
 DB 920 TCCCTGAGCAGCGCTGGGAGCGACCTTCTCTTGACTGCAACCACTATCAAAAGC 979  
 QY 1066 GTGCGGCAATGTGCGAGTGTGAGCTGCGACAGAGAGCGCGGAGCCAGCTATAGACATG 1125  
 DB 980 ATTTGGAAGCTTCCAAAGGCTGCGACCTGGGGAAGACAGACAGGCTTTGAGAACATG 1039  
 QY 1126 GAGCAATTTCTGTGAGCGTGAAGCAGTGGCGGAAATGTTGCGAATCGACTTCCAG 1185  
 DB 1040 GACTTCTTACTTGTGATTTAGAGCATATGACAGCGCCCTGAGAGCCATTTGACTTTGGG 1099  
 QY 1186 CCGGCGCGCGCTGGGATGTAAGAGATGACGACATGCTTTGGATGGGAGAGGSCAAT 1245  
 DB 1100 ACAGATGAGGAAAGAGAAATTCATTTGAAGAAAGATCAGAAAGAGAAAGATCCACA 1159  
 QY 1246 GCGGGCTGGAGAGAGAGCGGCTGAGCTGCGACATGCGCAAGG 1283  
 DB 1160 GAAAGGAGAGAGAGAGACACCACTTAAGAGACTGTGATG 1197

# RESULT 8

ABA99063  
 ID ABA99063 standard; DNA; 1597 BP.

XX ABA99063;

DT 15-JUL-2002 (first entry)

XX Murine muscle ring finger protein 3 (MURF-3) coding sequence.

XX Muscle ring finger; MURF-3; mouse; cardiant; microtubule;

KW Intermediate filament; striated muscle; cardiac hypertrophy;

KW heart disease; gene; ds.

XX Mus musculus.

OS Mus musculus.

FT Key Location/Qualifiers

FT CDS 299..1330

FT /\*tag= a

FT /product= "MURF-3"

PA (TEXTA ) UNIV TEXAS SYSTEM.

PI Olson EN, Spencer JA;

DR MPI: 2002-241506/29.

DR P-PSDB: ABB08277.

PT Novel muscle ring finger protein useful for drug screening, and for

PT diagnosing and treating diseases, particularly cardiomyopathies

XX Claim 4: Page 131-133; 134pp; English.

CC The sequence encodes murine muscle ring finger protein 3 (MURF-3). The invention relates to a purified muscle ring finger (MURF) protein.

CC selected from MURF-1, MURF-2 and MURF-3. The MURF proteins of the  
 CC invention are involved in microtubule and intermediate filament  
 CC stabilisation of striated muscle cells and have cardiact activity.  
 CC The MURF proteins are useful for screening a candidate substance  
 CC for MURF protein-binding activity, in a cell, cell-free system or  
 CC in vivo, and its effect on interaction of MURF with microtubules,  
 CC homodimerisation of MURF, MURF-1, MURF-2 or MURF-3 stabilisation  
 CC of microtubules, interaction of MURF with intermediate filaments,  
 CC e.g. desmin, vimentin and cytokeratin, and heterodimerisation of MURF.  
 CC The screened compounds are useful for treating and preventing of  
 CC cardiac hypertrophy and heart diseases. MURF proteins are useful as  
 CC antigens to immunise animals for the production of antibodies.

CC Sequence 1597 BP; 405 A; 404 C; 438 G; 350 T; 0 other;

Query Match 33.9%; Score 484.6; DB 24; Length 1597;  
 Best Local Similarity 69.2%; Pred. No. 1e-98;  
 Matches 677; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

QY 226 CTGCTAGGAGATCGCAACAACATGAGACAACTTGGAGAGAGAGCTCATTTGCCCATCTGC 285  
 DB 317 CTGATTCCTGATGGAACCGTATGAGAACCTGAGAGAGAGCTGATCTGCCCATCTGC 376  
 QY 286 CTGAGAGATGTTCTCAAGCCCGTGTGATCTTGCCTGCCAACAACCTGTGCCGCAAG 345  
 DB 377 CTGAGAGATGTTTACCAAGCCTGTGTCTATCTGCTGCCAACAACCTGTGCCGCAAG 436  
 QY 346 TGTGCCAAGAGACTTCTCCAGGCTCTAATCTCTGTGGCAATCCCGGGGCTCCACACG 405  
 DB 437 TGTGCCAAGAGACTTCTCCAGGCTCTAATCTCTGTGGCAATCCCGGGGCTCCACAG 496  
 QY 406 GTGTCTTCAAGAGAGAGATTCCTGATGCCATCTTGTAGGAGAGAGTGTCTGAGACAG 465  
 DB 497 TCCATGTCTGAGAGTGTGTTCCGTGCTGCCCTCGTCCGCAATGATGATGAGACCGG 556  
 QY 466 CATGTGTCTATGAGCTTGCAGCGGACCTGCTAGTGAAGACATCATTTGACATCTACAG 525  
 DB 557 CACGGGCTGTACCGCTGTGAGAGAACCTGTGTGAGAAACATCATTTGACATCTACAG 616  
 QY 526 CAGGAGCTCCCGGCACTGACGCGGCAAGGCTGACACGACCTCATCTGTGAGAGACG 585  
 DB 617 CAGGAGTGTCTCCAG---TCGGCCCTGCAGAAAGGACGCCACCGATGTGCAAGAGAC 673  
 QY 586 GAGGAGAGAGATCAACATCTACTGCTGAGTGCAGAGTGCACACCTGCTCTCTGTC 645  
 DB 674 GAGGAGAGAGATCAACATCTACTGCTGAGTGCAGAGTGCACACCTGCTCTCTGTC 733  
 QY 646 AAGGTTTCCGCGCCCAAGAGAGTGTGAGTGCACCTTGCACCATTTTACAAGC 705  
 DB 734 AAGGTTTCCGCGCTCACAGGCTGTGAGTGCACCTTGCACCATTTTACAAGC 793  
 QY 706 CAGAGAGTGTGAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 765  
 DB 794 CAGAGAGTGTGAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 853  
 QY 766 GCAGTATCACCCAGATGAGAGAGTGTGCCAGCATTTGAGACACAGCCGACAGAG 825  
 DB 854 AGCATATCTCTCAGCTGAGAGACTGTGAGAGTGCAGAGAGATGACACACAGAGTG 913  
 QY 826 AAGCACTGTAAACAGAGTGTGAGACCTGTGCGGGTTTGGAGAGAGAGAGAGC 885  
 DB 914 AAGGAGAGTGTGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 973  
 QY 886 GAGTGTGTTAAGAGCTGAGCGCGGAGAGAGAGAGTGTGAGCGGTGCGGCTTC 945  
 DB 974 GAGTGTGTTAAGAGCTGAGCGCGGAGAGAGAGAGTGTGAGCGGTGCGGCTTC 1033  
 QY 946 ATCCGCGAGTACGAGACCACTTGGAGGGCTCTCAAGCTGTGTGAGTCCGCCATCCAG 1005  
 DB 1034 ATCTCTCAGTACGAGAGAGAGCTGGAAGAGTCCACAGAGCTTGTGAGAGACCGCATCCAG 1093  
 QY 1006 TCCATGAGAGAGCGCGAGATGCTCTCTACTCCAGAGCGCAAGAGAGTGTATCAACAG 1065

DB 1094 TCCCTGATGAGCCCGGAGGGGCTACTTCTCTCAATGTCACAGAGCTCATCAAGAGC 1153  
 QY 1066 GTCCGGGCAATTCGAGAGTGTGAGAGAGCGCCGAGCCAGAGCTATGAGAGCATG 1125  
 DB 1154 ATTGTAGAGAGCTCCAGAGGCTGCGCAGCTGGGGAGACAGAGAGCTTTGAGACATG 1213  
 QY 1126 GAGCAATTCCTGTGAGCGTGTGAGACAGCTGGCCGAATTTGCGCAACATCGACTTCAG 1185  
 DB 1214 GACTACTTACTCTGTGAGCTTACAGACATAGCAGAGGCGCTTGGAGGCGCAATTCCTTGGG 1273  
 QY 1186 CCGGCGCGCGCTGGGAGATG 1204  
 DB 1274 ACAGGTAAAGATGTGATG 1292

RESULT 9  
 ABA83058  
 ID ABA83058 standard; DNA; 1764 BP.

XX ABA83058;

XX 05-FEB-2002 (first entry)

XX Human transcription factor TRFX-85 coding sequence.

DE Human; transcription factor; TRFX; cell proliferative disease;  
 KW autoimmune disease; inflammation; neurological disease;  
 KW developmental disorder; cancer; AIDS; infection; cytostatic; anti-HIV;  
 KW neuroprotective; antiinflammatory; gene therapy; ds.

XX Homo sapiens.

XX WO20012777-A2.

XX 04-OCT-2001.

XX 13-MAR-2001; 2001WO-US08117.

XX 13-MAR-2000; 2000US-0188986.

XX (INCY-) INCYTE GENOMICS INC.

XX Hillman JL, Baughn MR, Yue H, Lal P, Lu DM, Patterson C;  
 PI Azimzai Y, Bandman O, Tang YT, Mathur P, Shan P, Au-Young J;  
 PI Reddy R;

DR WPI: 2001-570896/64.  
 P-PSDB: ABB50234.

PT Novel transcription factor polypeptides, used to treat diseases  
 PT associated with altered activity and expression of TRFX, and to screen  
 PT for agents capable of modulating its activity -

PS Claim 11; Page 311; 327pp; English.

XX The present sequence is the coding sequence for a human transcription  
 CC factor. The transcription factor and its coding sequence are useful in  
 CC the diagnosis, treatment and prevention of diseases associated with  
 CC altered expression of the transcription factor e.g. cell proliferative,  
 CC autoimmune/inflammatory, neurological and developmental disorders. A  
 CC number of specific disorders/diseases are given in the specification,  
 CC including: arteriosclerosis, cirrhosis, hepatitis, cancers, AIDS,  
 CC allergies, anaemia, asthma, autoimmune thyroiditis, bronchitis, atopic  
 CC dermatitis, diabetes mellitus, emphysema, Goodpasture's syndrome, gout,  
 CC Grave's disease, multiple sclerosis, osteoarthritis, pancreatitis,  
 CC psoriasis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative  
 CC colitis, warts, Alzheimer's disease, Huntington's disease, Parkinson's  
 CC disease, stroke, and viral, bacterial, fungal and protozoal infections.

CC Sequence 1764 BP; 473 A; 399 C; 516 G; 376 T; 0 other;

Query Match 33.7%; Score 482; DB 22; Length 1764;  
 Best Local Similarity 66.6%; Pred. No. 4.1e-98;

Matches 705; Conservative 0; Mismatches 350; Indels 3; Gaps 1;

```

OY 226 CTGCTAGGGGATGGCAACAATGAGCAATGAGAGAGACCTATTTGCCCATCTGC 285
    ||| | -||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 141 CTGATCCAGAGATGGAAATCCCATGAGAACTTGAGAGAGAGCTGATCTGCCATCTGC 200
OY 286 CTGAGAGATTTCTCAAGCCCGGTGATCTTGCCTCCACACACACACTGTGCCGCAAG 345
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 201 CTGAGATGTTTACCAAGCAGTGTGATCTTGCCTGCCAGCACAACCTGTGCCGGAAG 260
OY 346 TGTGCCAAGAGAGCTTCCAGGCTCTAATCTCTGTGTGCAATCCGGGGCTCCACAAG 405
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 261 TGTGCTCAATGATCTTCCAGGCTGCAAACTCTACTGACCAAGCCGGGCTGAGTG 320
OY 406 GTGTCTTCAAGAGAGCTTCCAGTCCCATCTTGTAGAGACGAGTTGTCTGACAGAG 465
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 321 TCCATGCTCGAGAGGCGGTTTCCGCTCCACCTGCCGCCACGAGGTGATCATGTGATCGT 380
OY 466 CATGCTGTATGAGCTTCGACGGGAACCTGTAGTGGAGAAATCATTTGACATCTACAG 525
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 381 CACGAGTGTACGGCTTCGAGAGAACTGTGTGAGAAATCATCTGACATCTACAA 440
OY 526 CAGGAGTCTCCCGGCTGACGACGGCAAGGCTGAACAGACCTCATGTGTAGAGACAC 585
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 441 CAGGAGTCTCCAGTCCGCTGCAAGAGGCGAGT--CACCCTGTCAGAGAGACAC 497
OY 586 GAGGACGAGAAATCAACATCTACTGCTGAGCTGCGAGGTGCCACCTGCTCTCTGTC 645
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 498 GAAAGATGAGAAATCAACATCTACTGCTGAGCTGAGGTGCCACCTGCTCTCTGTC 557
OY 646 AAGGTTTTCGCGCCCAAGAGACTGTAGTGGCCCTCTCTGCCACCATTTACAAAGC 705
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 558 AAGGTTTTCGCGCCCAAGAGACTGTAGTGGCCCTCTCTGCCACCATTTACAAAGC 617
OY 706 CAGAGAGTGTAGTGTGAGAGTGTGATGCTGTGCTGGCGGCAATGACCTGTGTGAG 765
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 618 CAAGAAGTGTGAATCAATTAATCTGATCTGATGCTGTGTGCGGGGAGATGACCTGTGTCAG 677
OY 766 GCAGTGTACCCAGATGAGAGAGAGTGTGCCAGACCATTTGAGAGACAGCCGACAGAC 825
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 678 ACGATCATCTACCTGACGTGGAGATTCCTGAGTGCAGTACCAAGAGAACTCATCAGAGTA 737
OY 826 AAGCACTGTTAACCAAGAGTTCAGAGACCTGTGTGCGGCTTTGAGAGACGCAAGGCG 885
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 738 AAGGAGAGCTGAGAGCAAGATTTGACAGCTGTATGCCATCTGATGAGAGAAAGT 797
OY 886 GAACGCTTCAAGCACTGTGCGCGGAGGAGAGAGAGTGTGAGCGGCGTGGCGGCGCTTC 945
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 798 GAGTTGCTGCAAGCGGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 857
OY 946 ATCCGCGCAGTACGAGACCACTTGTGAGGCGCTCTCAAAAGCTGTGTGAGTCCGCAATCCAG 1005
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 858 ATCCAGCAGTACGAGACCACTTGTGAGGCGCTCTCAAAAGCTGTGTGAGTCCGCAATCCAG 917
OY 1006 TCCATGAGAGAGCCGCAATGCTCTTCACTCCAGCAGGCAAGAGAGTGTATCAACAAG 1065
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 918 TCCCTGAGAGAGCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 977
OY 1066 GTCCGCGCAGTACGAGACCACTTGTGAGGCGCTCTCAAAAGCTGTGTGAGTCCGCAATCCAG 1125
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 978 ATTTGTGAAGAGCTTCCAGAGGCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1037
OY 1126 GAGCAATTTCTGTGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1185
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 1038 GACTTCTTTACTTTGAGATTTAGAGACATFAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1097
OY 1186 CCGGCGCGCGCTGGGAGATGAAGAGAGATGACATGCTTTGTAGGAGAGAGAGAGAGAG 1245
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 1098 ACAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1157
OY 1246 GCGGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1283
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 1158 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1195

```

```

RESULT 10
AAH90037
ID AAH90037 standard; cDNA; 1231 BP.
XX
XX AAH90037;
D7 01-OCT-2001 (first entry)
XX
XX Human bone marrow cDNA, SEQ ID NO: 281.
DE
XX Human: bone marrow; antiinflammatory; cytosolic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200153453-A2.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 23-DEC-2000; 2000WO-US34960.
PF
XX
XX 21-JAN-2000; 2000US-0488725.
PR
XX 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250383.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
PI Ren F, Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Drmanac RT;
XX
XX WPI; 2001-488707/53.
DR
XX P-PSDB; AAM00918.
PT
XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
XX for treating e.g. cancer and immune deficiency disorders -
XX
XX Claim 1; Page 397; 648bp; English.
XX
XX The present sequence is one of 251 novel human polynucleotides
XX expressed in the bone marrow. The polynucleotide and the
XX polypeptide encoded by it are useful in the treatment of various
XX immune deficiencies and disorders. The deficiencies and disorders may
XX be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
XX infection, or may result from an autoimmune disorder, a coagulation
XX disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
XX suppression of an inflammatory response or treatment of a nervous
XX system disorder such as Alzheimer's disease. Detection of the presence
XX or increased expression of the polynucleotide or the protein it
XX encodes is useful for the diagnosis and/or prognosis of one
XX or more types of cancer. The polynucleotide and polypeptide can be
XX used as nutritional sources or supplements and in the screening of
XX chemical compounds as potential drugs.
XX
XX Sequence 1231 BP; 332 A; 305 C; 361 G; 233 T; 0 other;
SQ

```

Query Match 33.6%; Score 480.8; DB 22; Length 1231;  
Best local Similarity 67.2%; Pred. No. 6.9e-98;  
Matches 696; Conservative 0; Mismatches 337; Indels 3; Gaps 1;

```

OY 226 CTGCTAGGGGATGGCAACAATGAGCAATGAGAGAGACCTATTTGCCCATCTGC 285
    ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 167 CTGATCCAGAGATGGAAATCCCATGAGAACTTGAGAGAGAGCTGATCTGCCATCTGC 226

```

QY	286	CTGAGATGTTCTCCAAACCCGCGTGATCTTGACCTGGCCAAACAAACCGTGCAGG	34.5
Db	227	CTGGAGATGTTTACCAGGCACTGGTGCATCTTGGCGTGCAGCAACCTGTCCGGAG	286
QY	346	TGTGCCAAGCAGCGTCTTCCAGGCGCTCTATACCTCTGTGTGGCAATCCCGGGGCTCCACAGC	405
Db	287	TGTGCCAATGACATCTTCCAGGCTGCATACTCCCTACTGTGACCAACCGGGGCACTCAGTG	346
QY	406	GTCGTTTCAGGAGCAGCTTTCGGATGCCATCTTGTAGCCAGAGGTTGTCTGTGAGG	465
Db	347	TCCATGTCGTGGAGGCGGTTTCCGCTGGCCCACTGGCCGACAGGTGATCATGATCGT	406
QY	466	CATGGTGTCTATGGGCTGCAGGCGAACCGTATGTAGAGCAACATCTTGCATCTCAAG	525
Db	407	CACGAGGTGTACGGCCTGCAGAGGAACCTGCTGTGTGGAGAACATCTCATCTACATA	466
QY	526	CAGGAGTCTCCCGGCCACTGCGACGCAAGGCTGAACACAGCACTCATGTGTGAGGAGCAC	585
Db	467	CAGGAGTCTCCAGTCGGGCGCGCTGCAGAG--GGCAGTCACCCCATGTGCAGAGCAC	523
QY	586	GAGGACGAGAGATCAACATCTACTGTCCCTGAGCTGCGAGGTGGCCACCTCTCTCTGC	645
Db	524	GAGGTGTGAGAAATCAACATCTACTGTCTACAGTGTGAGGTGCCACCTGCATCTATGC	583
QY	646	AAGGTTTTCCGGGCCCAACAGACAGTGTAGGTGGGCCCTCTGCCACATTTACAAGCG	705
Db	584	AAGGTGTTTGGATCCACAAGGCTTCGAGGTGGCCCATGTGCAGAGTGTCTTCCAGGGA	643
QY	706	CAGAGAGTGAGCTGAGCGATGGCATCGGATGTCTGTGGCGGGCAATGACCGTGTGAG	765
Db	644	CAAAAGACTGAATCAATATGTAATCTTCATGCTGTGGTGGCGGGAATGACCGTGTGAG	703
QY	766	GCAGATATCACCACATGAGGAGGAGTGTCCCAAGCATTTAGGAGACAACACCGCAGACG	825
Db	704	ACCATCATCTACCTACGCTGGAGGATTCGCCCTCGAGTGAACCAAGGAGAACTCATCAGGTA	763
QY	826	AAGCAACTGTTTAAACCGAGAGTGTGAGACCTGTGGCGGTTTTTGGAGGACGCAAGGGC	885
Db	764	AAGGAAGAGCTGAGCGAAGATTTGACAGGTTGTATGCCATCTGGATGAGAAGAAAGT	823
QY	886	GACATGCTTCAAGCACTGGCCCGGGAGCAAGAGAGAAATTGCAAGCGCGTGGCGGGCCTC	945
Db	824	GAGTGTGCTCACCGGATCTACGCGAGGAGGAGAAAAAGCTTGTAGCTTATCAGAGGCGCTC	883
QY	946	ATCCGCCAGTACGAGAACCACTTTGGAGGGCTCTCAAGCTGGTGGAGTCCGCCATCCAG	1005
Db	884	ATCCAGCACTACCAAGAGAGCACTGGCAAACTCCACAAGTAAGTGTGGAACTGGCCATCCAG	943
QY	1006	TCCATGAGAGAGCCGAGATGGCTCTCTATCTCCCTCCAGCAGGCAAAAGAGCTGATCAACAG	1065
Db	944	TCCCTGAGAGACCTGTGGGGAGCCACTTCTCTTGTGACGTCCCAAGCAATCATCAAAAGC	1003
QY	1066	GTCGGGGCAATGTCANAAGTGGAGCTGGCAGAGACGGCCGAGCAAGCCATGTAGAGCATG	1125
Db	1004	ATTGTGGAAGCTTCCAAAGGCTGCCAGCTGGGGGAAGACAAGAGGGCTTTGAAGAAATG	1063
QY	1126	GAGCAATTTCTGTGAGCGTGGAGCAAGTGGCCGAATGTTTGGCAACCTATGCATTTCCAG	1185
Db	1064	GACTTCTTACTTTTGTGATTTAGAGCATATGACAGACGCCCTTGAGAGGCATTTGACTTTGGG	1123
QY	1186	CCGGGCGCGCTGGGGATGAAGAAGATGACAGACATGGCTTTTGGATGGGAGGAGGGCAAT	1245
Db	1124	ACAGATGAGAGAGAGAAATTCATTTGAAGAAGAAGATCAGGAAGAGAAAGTCCACA	1183
QY	1246	GGGGGCGCTGGAGAGG 1261	
Db	1184	GAGGGAGAGGAAG 1199	





Db 428 AAAAATTG-----ACGACCCCATGTGTGAAGACATGAAGAGGAAAGCATCATC 481  
QY 607 TACTGCTGAGCTGCGAGGTGCGCCACCTCTCTCTGCAAGTTTGGCGCCACAAG 666  
Db 482 TATGTCTGAAGTGTGAAGTGGCCACCTGCTCTTGTGCAAGTTTGGCGCCACAAG 541  
QY 667 GAGTGTGAGTGGCCCTGCTGCGCCACCTTACAAAGGCAAGAGTGAAGTGGAGAT 726  
Db 542 GACTGCGCAGGTGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601  
QY 727 GGCATCGGATGCTGTGTGGCGGCAATGACCGTGTGCAAGGCAAGTGAAGTGAAG 786  
Db 602 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 661  
QY 787 GAGTGTGCGCAGTGTGAGGACAAACAGCGGCAAGAGCAAGTGTGAAGCAAGG 846  
Db 662 GACACCTGTAAACATTTGAGAGTGTGCTGCAAAAGCAAGCAAGCAAGTGTGAGAA 721  
QY 847 TTGAGACCTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906  
Db 722 TTTGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 781  
QY 907 CGGAGCAGAGAGAGAGAGTGTGAGCGCGGCGGCGCTCATCCGCGCAGTACGAGAC 966  
Db 782 CGAACACAGAGAGAGAGAGTGTGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 841  
QY 967 TTGAGAGGCTCTCAAGTGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026  
Db 842 CTGGAGAACGTAATCCAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 901  
QY 1027 GCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1086  
Db 902 GCAGATTTCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961  
QY 1087 GAGCTGCGAGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1146  
Db 962 TTTGAGATGAGAGAACTGAGAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1021  
QY 1147 GAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1206  
Db 1022 AATGAG 1081  
QY 1207 GAGGATGAG 1266  
Db 1082 GATGAG 1141  
QY 1267 CTGACGTGCCAGAGGCTC 1286  
Db 1142 GAAATGTTCAATATGATC 1161

RESULT 13  
AAS25855  
ID AAS25855 standard; cDNA, 1762 BP.  
XX AAS25855;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
XX  
XX Human cDNA encoding a novel secreted protein, Seq ID 34.  
XX  
XX Human: immunosuppressive; antiarthritic; ss; antirheumatic;  
KW cytoprotective; cardiant; vasotropic; cerebroprotective; noctropic;  
KW neuroprotective; antibacterial; virucide; fungicide; ophtalmological;  
KW vulnerary; secreted protein; rheumatoid arthritis;  
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;  
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;  
KW corneal infection; wound healing; epithelial cell proliferation;  
KW skin ageing; food additive; preservative; antiproliferative.  
XX Homo sapiens.  
XX  
XX

PN W020015322-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01341.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214686.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234597.  
PR 25-SEP-2000; 2000US-0234598.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235836.





```

Db 619 AGACTGCCAGTGCGCTCCCTCCTCATCATGTGTTCACAGACAGAAAGTCTGAGTCACTGA 678
QY 726 TGGCATCCGAGTCCGAGTGGGCGGCGCAATACCGTGTGCAGGCGATGATCAACCCAGATGGA 785
Db 679 TGGCATCCGAGTCCGAGTGGGCGGCGCAATACCGTGTGCAGGCGATGATCAACCCAGATGGA 738
QY 786 GGAGGTGGCCAGACCATTTGAGGACAAACCCGACAGACAGACAGACATGTTAAACAGAG 845
Db 739 AGACACCTGCAAAATATTCAGAGGAATGTTGCAGAAAACAAACAAAGACGTTTGTGAA 798
QY 846 GTTCGAGACCTGTGCGCGGTTTGGAGAGCGCAAGGCGCAACGCTTCAAGCATGGC 905
Db 799 GTTGATTTACCTGTATGCGATTTTGGAGAGAGAAATGAATGACCAAGTCATTTAC 858
QY 906 CCGGAGAGAGAGAGAAATGTCAGCGCGCTGCGGCGCTCATCCGCGATAGAGAGACCA 965
Db 859 CCGAAGCCAGAGGAGAAACTGGAACTGTCTGCTGTATCAAAAAGTATTTCTGATCA 918
QY 966 CTTGGAGGCTCCTCAAAAGCTGTGAGTCCGCGCATTCACATCCAGAGAGCGCGAGAT 1025
Db 919 TTTGAGAGAACGTCCTCAAAATGTTGATGAGTCAAGATTTCAATTTATGATGAGCCAGAAAT 978
QY 1026 GGCCTCTTACTTCCAGACGCAAAAGAGCTGATCAACAGCTGCGGCAATGTGAAAGT 1085
Db 979 GCGAGTGTCTTTCAGAAATGCCAAACCTGCTAAAAAATATCTCAAGAGCATCAAGGCG 1038
QY 1086 GGAGGTGCGAGAGCGCGGCGGAGCGGAGTATGAGCATGAGAGATTTCTCTGTGAGCGT 1145
Db 1039 ATTTGATGATGGGAAATATGAAACATGCTATGAGAACTATACACTTCCAGAGTCAACT 1098
QY 1146 GGAGCAGCTGCGCGGAAATGTTGCGAATCGACTTCAGCGCGGCGCGCGTGGGAGTA 1205
Db 1099 CAATGAGAGAGAAAGATTAATACGTGAATTTGACTTTTACAGAGAAATGAAGATGAAGA 1158
QY 1206 AGAGATATACGATGCGTTTGGATGGGAGAGGCGAATGCGGCGCTGGAGAGAGAGCG 1265
Db 1159 AGAAGAAAGAGGCG--GGAGAGAGAGAAAGAGAGAGAAAGTGGGAGAGAGAAAGCA 1216
QY 1266 GCTGGA 1271
Db 1217 GTAGAA 1222

RESULT 14
AAH90117 standard; cDNA; 1183 BP.
XX AAH90117;
AC 01-OCT-2001 (first entry)
XX 01-OCT-2001 (first entry)
DE Human bone marrow cDNA, SEQ ID NO: 474.
XX Human; bone marrow; antiinflammatory; cytosolic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
XX Homo sapiens.
OS Homo sapiens.
XX WO200153453-A2.
PN 26-JUL-2001.
XX 26-JUL-2001.
XX 23-DEC-2000; 2000MO-US34960.
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.

```

```

PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250583.
XX (HYSE-) HYSEQ INC.
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Drmanac RT;
XX WPI: 2001-488707/53.
DR P-PSDB: AAM00998.
PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT for treating e.g. cancer and immune deficiency disorders -
XX Claim 1: Page 608-609; 648pp; English.
XX PS
XX CC The present sequence is one of 251 novel human polynucleotides
XX CC expressed in the bone marrow. The polynucleotide and the
XX CC polypeptide encoded by it are useful in the treatment of various
XX CC immune deficiencies and disorders. The deficiencies and disorders may
XX CC be genetic, or may be caused by a viral (e.g. HIV), bacterial or fungal
XX CC infection, or may result from an autoimmune disorder, a coagulation
XX CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
XX CC suppression of an inflammatory response or treatment of a nervous
XX CC system disorder such as Alzheimer's disease. Detection of the presence
XX CC of increased expression of the polynucleotide or the protein it
XX CC encodes is useful for the diagnosis and/or prognosis of one
XX CC or more types of cancer. The polynucleotide and polypeptide can be
XX CC used as nutritional sources or supplements and in the screening of
XX CC chemical compounds as potential drugs.
SQ Sequence 1183 BP; 294 A; 318 C; 327 G; 244 T; 0 other;

Query Match 29.3%; Score 419; DB 22; Length 1183;
Best Local Similarity 69.1%; Pred. No. 4.4e-84;
Matches 588; Conservative 0; Mismatches 260; Indels 3; Gaps 1;

QY 226 CTGCTAGGAGATGCGCGCAACATGAGACACTTGGAGAGAGAGCTGATTTGCCCATCTGC 285
Db 175 CTGATCCAGATGCGGATGCCATGCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 234
QY 286 CTGAGATGTTCTCCAGACCGCGGAGTGTCTGCGCTGCCAAGACACACCTGTGCCGAG 345
Db 235 CTGAGATGTTCTACAGACCGAGTGTCTGCGCTGCCAAGACACACCTGTGCCGAG 294
QY 346 TGTGCCAAGACAGCTTCCAGCGCGCTTAATCTGTGTGCGCAATCCCGGGGCTCCAGAG 405
Db 295 TGTGCCAATGACATCTTCCAGCGCGTGAATCTCTACTGTGAGACCGCGGAGAGAGAG 354
QY 406 GTGTCTTCCAGAGAGAGCTTCCAGATGCCATCTGTAGAGAGAGAGAGAGAGAGAGAG 465
Db 355 TCCATGTCTGAGGCGCGTTTCCGCTGCCCGCCAGCGCGCGAGAGAGAGAGAGAGAG 414
QY 466 CATGTGTCTATGCGCTGACAGCGGAGACCTGTAGTGAAGAAATCATTTGACATTCACAG 525
Db 415 CAGGAGGTGTACGCGCTGACAGAGAGAGACCTGTGAGGAGAAATCATGACATTCACAAA 474
QY 526 CAGGAGTCTCCCGCGCGACACGCGCAAGGCTGAACGACGACGCTATGTGAGAGAGAG 585
Db 475 CAGGAGTCTCCAGAGTGGCGCGCTGACAGAG--GGCACTACCGCATGTGCAAGAGAGAG 531
QY 586 GAGGAGAGAGAGATCAACATCTACTGCTGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAG 645
Db 532 GAGATGAGAGAAATCAACATCTACTGTCTACGTGTGAGAGTGCAGAGTGCAGAGTGC 591
QY 646 AAGCTTTGCGCGCGCAAGAGAGTGTGAGTGGCGCTGTGCCACCATTTTCAAGAGCG 705
Db 592 AAGGTGTTGGAGATCCAAAGGCTGCGAGGTGGCGCCCATGTCAGAGTCTTCCAGAGGA 651
QY 706 CAGAAGAGTACGTGAGAGAGATGATGCGGAGTGGGAGAGAGAGAGAGAGAGAGAGAG 765
Db 652 CAAAAGACTGAAGTATCTGATCTCCATGCTGTGAGGCGGAGAGAGAGAGAGAGAGAGAG 711

```









to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain cell proliferation, to transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. The present sequence encodes a novel secreted protein of the invention.

Query Match 16.0%; Score 229.2; DB 22; Length 587;  
Best Local Similarity 70.9%; Pred. No. 9e-42;  
Matches 333; Conservative 1; Mismatches 129; Indels 7; Gaps 2;

QY 187 GAGCGCGCGGGGAGTCAAGCGGTTCAGCGCGCTGCTAGGCGATGCGCACAC 246  
DB 124 GACAGCGAGGAGATGAGCCATCTTGATTAACAAATCTTTCCAAAGAGCAGAAC 183  
QY 247 ATGACAACTTGAGAGAGAGCATTTGGCCCATCTGCTGAGATGTTCTCCAGCCC 306  
DB 184 ATGGATTAATGAGAGAGCACTATCTGCTGCTAGAGATGTTACAGAACT 243  
QY 307 GGGGATATTCGCCCTCCACAGCAACCTGCGCGAGGTGCGACAGCTTCCAG 366  
DB 244 GGGGATATTCGCCCTCCACAGCAACCTGCGCGAGGTGCGACAGCTTCCAG 303  
QY 367 GCGTCAATCTCTGTGGCAATCCCGGGCTCCACAGGCTGCTTCCAGAGAGAG 426  
DB 304 GCGTCAATCCGATTTGGCCCAAGAGAGGTAACCCATGCGATCAGGGGCGCAT 363  
QY 427 CGATGCCATCTTTGAGCAGAGGTTGCTCTGAGCAGGATGTCATAGCTGAG 486  
DB 364 CGTGCCCATCTCTGTAGCAACCTGTGAGAAATGTCAGTATTTTCAG 423  
QY 487 CGGAACCTGTAGTGGAGCATGATGATCAGAGAGAGAGAGAGAGAGAGAG 546  
DB 424 AGGAACCTGTGTGAGAAATATCATTGATCAGAGAGAGAGAGAGAGAGAG 480  
QY 547 CAGCGCAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606  
DB 481 ---GAAAGAGAAATCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 537  
QY 607 TACTGCTGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 656  
DB 538 TACTGCTGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 586

RESULT 19  
AAS26303  
ID AAS26303 standard; cDNA; 446 BP.  
XX  
AC AAS26303;  
XX  
DT 07-NOV-2001 (first entry)  
DE Human cDNA encoding a novel secreted protein, Seq ID 482.  
XX  
XX Human: immunosuppressive; antiarthritic; ss; antirheumatic;  
KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;  
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
KW vulnary; secreted protein; rheumatoid arthritis;  
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;  
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;

KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;  
KW corneal infection; wound healing; epithelial cell proliferation;  
KW skin ageing; food additive; preservative; antiproliferative.  
OS Homo sapiens.  
XX  
PN WO200155322-A2.  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01341.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225256.  
PR 14-AUG-2000; 2000US-0225257.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 01-NOV-2000; 2000US-0241826.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249267.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 01-DEC-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251889.  
 PR 08-DEC-2000; 2000US-0251990.

PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI  
 XX WPI: 2001-488783/53.  
 DR P-PSDB; AAU16316.  
 XX  
 XX New nucleic acid molecules encoding 461 human secreted proteins for  
 PT diagnosing, preventing, treating or ameliorating medical conditions and  
 PT used as food additives or preservatives -  
 XX  
 PS Claim 1; SEQ ID No 482; 980bp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules and their  
 CC encoded secreted proteins. The nucleic acids and proteins are used to  
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They  
 CC are also used in diagnosing a pathological condition or susceptibility  
 CC to a pathological condition. Antibodies to the proteins can also  
 CC be used in alleviating symptoms associated with the disorders and in  
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
 CC immunosorbant assays (ELISA). Disorders which are diagnosed or treated  
 CC include autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi  
 CC and ocular disorders e.g. corneal infection, and many other  
 CC disorders listed in the specification. The polypeptides can also  
 CC be used to aid wound healing and epithelial cell proliferation, to  
 CC prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
 CC as a food additive or preservative to increase or decrease storage  
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
 CC minerals, cofactors and other nutritional components. The present  
 CC sequence encodes a novel secreted protein of the invention.

Query Match 11.3%; Score 162; DB 22; Length 446;  
 Best Local Similarity 72.5%; Pred. No. 8.5e-27;  
 Matches 222; Conservative 0; Mismatches 82; Indels 2; Gaps 1;

OY 226 CTGCTAGGGGATGGCGCACATGAGACACTTGAGAGACGCTCATTTGCCCATCTGC 285  
 ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 138 CTGATCCAGGATGGGATGGGATGAGAACTTGAGAGACGCTCATTTGCCCATCTGC 197  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 OY 286 CTGAGATGTTTCCAAAGCCCGTGATCTTGCCCTGCCCAACAACCTTGCCGCAAG 345  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 198 CTGAGATGTTTCCAAAGCCAGTGGTATCTTGCCCGTCCGCAACAACCTTGCCGCAAG 257  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 OY 346 TGTGCCAAGCAGCTTTCCAGGCGCTTAATCTCTGTGGCAATCCGGGGCTGCACAAG 405  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 258 TGTGCCAATGACATCTTCCAGGCTGCAAAATCCCTACTGACCAACCGGGGAGCTCAGTG 317  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 OY 406 GTGTCTTCAGAGAGAGCTTCCGATGCCATCTTGTAGGACAGAGGTGTCCTGGAAGG 465  
 || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 318 TCCATGCTGGAAGCCGTTTCCCTGCCCCACCTGCGCCACGAGNAGATCATGGAATCGT 377  
 || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 OY 466 CATGTTCTTATGGCCTG--CAGCGAACCCTGCTAGTGAAGAAATCATGATGACTTACA 523  
 || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 378 CACGAGGTGACGGCTGGCAANAAGACGTGGGTGGAAACATCATCATCATCTACA 437  
 || ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 OY 524 AGCAGG 529  
 || |||  
 Db 438 AACAGG 443

RESULT 20  
 AAC75760  
 ID AAC75760 standard; cDNA; 650 BP.







GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 02:23:10 ; Search time 60 Seconds  
(without alignments)  
7314.241 Million cell updates/sec

Title: US-09-908-988B-1  
Perfect score: 1431  
Sequence: 1 aagagagctgtagcagagagtgcttcc 1431

Scoring table:  
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgcn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgcn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgcn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
4: /cgcn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
5: /cgcn2\_6/ptodata/1/ina/PTUS.COMB.seq:\*  
6: /cgcn2\_6/ptodata/1/ina/Backfileseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	485.2	33.9	1757	US-09-484-970B-37	Sequence 37, Appl
2	70	4.9	3262	US-09-561-989-9	Sequence 9, Appl
3	66.6	4.7	1926	US-09-249-585A-2	Sequence 2, Appl
4	66.6	4.7	2580	US-09-050-863-2	Sequence 2, Appl
5	66.6	4.7	2580	US-09-359-081-2	Sequence 2, Appl
6	66.6	4.7	5452	US-09-130-114-1	Sequence 1, Appl
7	66.6	4.7	9600	US-08-910-647-1	Sequence 1, Appl
8	66.6	4.7	9600	US-09-620-925-1	Sequence 1, Appl
9	66.6	4.7	10596	US-07-884-811-15	Sequence 15, Appl
10	66.6	4.7	10596	US-08-087-783A-15	Sequence 15, Appl
11	66.6	4.7	10596	US-08-087-783A-15	Sequence 15, Appl
12	66.6	4.7	10596	US-08-194-088B-15	Sequence 15, Appl
13	66.6	4.7	10596	US-08-194-087-15	Sequence 15, Appl
14	66.6	4.7	10596	PCT-US93-04648-15	Sequence 15, Appl
15	64	4.5	1926	US-09-249-585A-4	Sequence 4, Appl
16	64	4.5	1926	US-09-130-114-2	Sequence 2, Appl
17	59.8	4.2	7218	US-08-232-463-14	Sequence 14, Appl
18	54.2	3.8	3489	US-08-728-323A-1	Sequence 1, Appl
19	54.2	3.8	3489	US-09-298-568-1	Sequence 1, Appl
20	54.2	3.8	32207	US-08-770-379-20	Sequence 20, Appl
21	54.2	3.8	32207	US-08-757-669A-20	Sequence 20, Appl
22	54.2	3.8	32207	US-09-230-371A-20	Sequence 20, Appl
23	53	3.7	2338	US-08-425-069-1	Sequence 1, Appl
24	52.4	3.7	2338	US-08-317-844B-1	Sequence 1, Appl
25	52.4	3.7	2793	US-08-209-747-1	Sequence 1, Appl
26	52.4	3.7	2793	US-08-458-298-1	Sequence 1, Appl
27	50.8	3.5	4411529	US-09-103-840A-1	Sequence 1, Appl

Considered  
11/11/05  
No.1

28	45.6	3.2	1505	US-07-915-246-1	Sequence 1, Appl
29	45.6	3.2	2214	US-08-864-038A-1	Sequence 1, Appl
30	45.6	3.2	3331	US-08-864-038A-2	Sequence 2, Appl
31	45.6	3.2	3331	US-08-864-038A-4	Sequence 4, Appl
32	45.6	3.2	9551	US-08-056-200-93	Sequence 93, Appl
33	45.6	3.2	9551	US-08-800-644-93	Sequence 93, Appl
34	44.6	3.1	1995	US-08-425-069-3	Sequence 3, Appl
35	44.6	3.1	1995	US-08-317-844B-3	Sequence 3, Appl
36	44.4	3.1	4214	US-09-123-135-1	Sequence 1, Appl
37	44.2	3.1	1767	US-07-903-466-2	Sequence 2, Appl
38	44.2	3.1	1767	PCT-US93-05794-2	Sequence 2, Appl
39	44.2	3.1	3018	US-07-903-466-1	Sequence 1, Appl
40	44.2	3.1	3018	PCT-US93-05794-1	Sequence 1, Appl
41	43.6	3.0	4852	US-07-853-913-3	Sequence 3, Appl
42	43	3.0	2830	US-09-010-928B-1	Sequence 1, Appl
43	43	3.0	3223	US-07-980-528-1	Sequence 1, Appl
44	43	3.0	51255	US-08-781-891-209	Sequence 209, App
45	42.8	3.0	765	US-09-603-785-524	Sequence 524, App

ALIGNMENTS

RESULT 1  
US-09-484-970B-37  
Sequence 37, Application US/09484970B  
Patent No. 6426186  
GENERAL INFORMATION:  
APPLICANT: Jones, Karen A.  
APPLICANT: Volkmut, Wayne  
TITLE OF INVENTION: BONE REMODELING GENES  
FILE REFERENCE: PB-0014 US  
CURRENT APPLICATION NUMBER: US/09/484,970B  
CURRENT FILING DATE: 2000-01-18  
NUMBER OF SEQ ID NOS: 177  
SOFTWARE: PERL Program  
SEQ ID NO 37  
LENGTH: 1757  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No. 6426186 247384.1CBI  
US-09-484-970B-37

Query Match	Similarity	Score	DB 4:	Length	1757:
Best local	66.8%	Pred. No. 6.2e-106;			
Matches 707;	Conservative	0;	Mismatches 348;	Indels 3;	Gaps 1;
QY	226	CTGCTAGGAGGATCGGCAACATGGACACTTGGAGAGAGAGCTATTTGCCCATCTGC	285		
DB	143	CTGATCCAGATGATGGATGGAATCCCATGGAGACTTGGAGAGAGAGCTATTTGCCCATCTGC	202		
QY	286	CTGGAAGATGTTCTTCCAAAGCCCTGCTATCTTGGCCCAACCAACCTGCGGCAAG	345		
DB	203	CTGGAATGTTTCCAAAGCCAGGAGTGTATCTTGGCCGAGCAACACTTGGCGGAG	262		
QY	346	TGTGCAACAGACGCTCTCCAGGCTTATCTTATCTTGGCAATCCGGGCTCCCAAG	405		
DB	263	TGTGCAATATATTTTCCAGGCTCTTATCCCGATTTTGGCCCAAGAGGATACACC	322		
QY	406	GTCCTTCCAGAGAGAGCTTCCGATCCCATCTTGAAGGAGAGAGTGTCTCGACAGG	465		
DB	323	ATGCGATCAGAGGAGGCGGATCCGCTGCGCATCTGTAGACATGAGAGTGTGATGA	382		
QY	466	CATGCTCTATGAGGCTTCCAGGCGGCAACCTGCTAGTGAAGACATCATGACATCAAG	525		
DB	383	CATGAGGTATATGACTTCAAGAGAACCTGCTGTGGAACATCATGACATCAAG	442		
QY	526	CAGAGTCTTCCCGGCGGCAACAGGCTGAACAGCACTTATGTGAGAGAGAG	585		
DB	443	CAGAGTCTTCCCGGCGGCGGCGGCTGCAAGAGGAGAGT---CACCCCATGTGCAAGAGAGAG	499		

QY	586	GAGGACAGAGATCAATCATCTACTGCGCTGAGCTCGAGGTGCGCCACCGTCTCTGCG	645
Db	500	GAGATGAGAAATCAACATCTACTGTCTCACGTGTGAGGTGCCACCTGTCTCATGTGCG	559
QY	646	AAGGTTTTGGGCGCCCAAGAGACTGTGAGGTGGCCCTCTGCGCCACATTTACAAGCG	705
Db	560	AAGGTGTTTTGGGATCCACAGAGCGCTGCGAGGTGGGCCCATTTGCAGAGTGTCTTCCAGGA	619
QY	706	CAGAAAGTGTGAGCTGAGCGATGGGATGCGGATGCGTGTGGGGGCAATGACCTGTGCGAG	765
Db	620	CAAAAGCTGTACTGGAATTAATGTATCTTCATGCTGTGGTGGGGAAATGACCGTGTGCGAG	679
QY	766	GCAGTGATCAACCCAGATGAGAGAGGTGTGCCAGACCATTTGGAGAGCAAGCCGCAAGAC	825
Db	680	ACCATCATCACTCAGCTGAGAGATTTCCGCTGAGTGCACCAAGGAAACAGTATACCAGTA	739
QY	826	AACCAACTGTAAACCCAGAGGTGTGAGACCCCTGTGCGCGGTTTTGGAGAGCGCAAGGCG	885
Db	740	AAGGAAAGCTGTAGCCCAAGAACTTTGACAGTTGATGCGATCTCGATGTAGAAAGAAAT	799
QY	886	GAACTGCTTCAAGCATCTAGGCCCGGAGACAGAGAGAGAGTGTGCACGGTGTGGGGCCCTC	945
Db	800	GAGTTGCTGACGGGATATCAGCAGAGAGCGSAGAAAAAGCTTAGTTATCGAGGCCCTC	859
QY	946	ATCCGCCAGTACGAGACCCACTTGTGAGGGCTCTCAAACTGTGTGAGTCCGCCATCCAG	1003
Db	860	ATCCAGCAGTACCAAGAGGAGCGACTGGACAAAGTCCAAAGCTGTGTGAAATCTCCATCCAG	919
QY	1006	TCCATGGAGAGCGCGCATGATGGCTCTTACTCCAGCAGGCAAAAGAGACTATCAACAAG	1063
Db	920	TCCCTGGACAGAGCTGGGGGAGCCACTCTCTCTTGTACTGCCAAACACTATCTCAAAAGC	979
QY	1066	GTCGGGGCAATGTCCAAAGTGTGAGCTGTGGCAGAGCGCCGGAAGCCAGCGCTATGAGACATG	1123
Db	980	ATTGTGTGAAGCTTCCAAAGGCGCTGCACTGTGGGGAAGACAGAGCGGCTTTGAGAACATG	1033
QY	1126	GAGCAATTTCTGTGAGCGGTGAGACAGTGGGCCGAATGTTGGCAACCATGTGACTTCCAG	1183
Db	1040	GACTTCTTACTTTGGATTTAGAGCACATATAGCMAGCCCGCTGAGAGCGCATTTGACTTTGGG	1093
QY	1186	CCGGGCGCCGCTGGGGATGAAGAGATGACACATGTGCTTTGGATGTGGGAGAGGGCAAT	1243
Db	1100	ACAGATGAGGAAGAGAGAAATTCATTTGAAGAGAAAGATCAGGAAGAGAGATGCCACA	1153
QY	1246	GCGGGCTGGAGAGAGAGCGGCTGTGACGTGCCCAAGG	1283
Db	1160	GAAGGAAAGAAAGAGACCACTGTAAGGAGCTGGATG	1197

```

RESULT 2
US-09-561-989-9
; Sequence 9, Application US/09561989
; Patent No. 6468750
; GENERAL INFORMATION:
; APPLICANT: KOLLER, Klaus-Peter
; TITLE OF INVENTION: No. 6468750el Cell Regulation Factor TTO20
; FILE REFERENCE: seq
; CURRENT APPLICATION NUMBER: US/09/561,989
; CURRENT FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3262
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-561-989-9

Query Match      4.9%; Score 70; DB 4; Length 3262;
Best Local Similarity 47.1%; Pred. No. 1.4e-07;
Matches 214; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

Yy 563 AGCAACCTCATGTGTAGAGACGACGAGACGAGACAGATCAACACTTACTGCGCTGACGCTGCG 622

```

Db	329	AGGTGTCACACCTGCACAGACGCCAGCGCTGGAGAAACCAAGCATGTACTGCGTCAATGCA	388
Qy	623	AGGTGCGCCACCGTCGCTCTGTGCAAGGTTTTCGGGGCCCAAAAGACTGTAGGTGGCCC	682
Db	389	AATATGCCCGGTGTGTACCAAGTGCCTTTGGAGGAGGGGCAAACTCCAGCCACCAAACTCAAGG	448
Qy	683	CTCTCCCAACCATTTACAAAGCCCGAGAGAGTGAAGCTTGAAGGATGGCATGCCATGCTGG	742
Db	449	CTCTGGGGGCCATGTGGAAACTACTTAAGAACCAAGCTTCCCAAGCGCTGTGAACGAGACTGT	508
Qy	743	TGGCGGGCAATATACCGTGTGCAGGCGCATGATCACCAAGATGGAGAGGTGACCGACACA	802
Db	509	CAGACAGAGGCCCAAAACAAACCCAAAGAGTTTCTGTGTACACGTGCGCAACATGCTCCAGACA	568
Qy	803	TTGAGGACACACGCCGCGACAGAGAAGCAACTGTTAAACAGAGGTTGAGACCCCTGTGCG	862
Db	569	TCCAGGAACACAGTGTGGAGTTTGAAGCTGTCTGTGTGGGCCCAATGTGATGCCCTCATCG	628
Qy	863	CGGTTTGGAGAGAGCGCACGAGGCGCACTGTTCAAGCAGCTAGTGGCCCGGAGACGAGAGAGA	922
Db	629	ATGCCCTCAACAGAAAGAAAACCCAGCGCTGGGCCGCGTCAACAAAGAGCAATGAGACACA	688
Qy	923	AGTTGCACAGCGGTGCGGGGCTCATCCGCCAGTACGAGAGACCATTGGAAGGCTCTCAA	982
Db	689	AGCTGAAGGTGGTTCGAATCAGATCTCTACTCCACAGTGAATTTGCCCAAGCACAG	748
Qy	983	AGCTGTGAGTCCGCCATCCAGTCCATGGAGGA	1016
Db	749	GTCCTCATGAGTACTGCTTGGAGGTGATTAAGGA	782

```

RESULT 3
US-09-249-585A-2
: Sequence 2, Application US/09249585A
: Patent No. 6417002
: GENERAL INFORMATION:
: APPLICANT: Horlick, Robert
: TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
: FILE REFERENCE: 0867/0D905
: CURRENT APPLICATION NUMBER: US/09/249,585A
: CURRENT FILING DATE: 1999-02-11
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 2
: LENGTH: 1926
: TYPE: DNA
: ORGANISM: Epstein Barr Virus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(1926)
: OTHER INFORMATION: coding strand of EBNA-1 DNA
: US-09-249-585A-2

```

[illegible]





```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-910-647-1

```

Query Match	4.78;	Score 66.6;	DB 4;	Length 9600;
Best Local Similarity	45.5%;	Pred NO 1	3e-06;	

	Matches	237;	Conservative	0;	Mismatches	284;	Indels	0;	Gaps
OY	760	GTGCAGGCACTGATCATCCACAGTGGAGAGAGTGTCGCCAAGACCATTTAGAGCAACAAGCGCC	819						
Db	719	GGGCGAGGCGACGAGAGGGGCGAGGAGCAGAGAGGAGGGGCGAGAGCGGCGAGAG	778						
OY	820	AGACAGAGCAACTTTAAACCAGAGGTTTCGAAGACCCCTGCGCCGGTTTTGGAGGAGCGC	879						
Db	779	GGGCGAGGAGGGGCGAGAGGGGCGAGGAGCACAGAGAGGGGCAGGACCAAGAGAGAGGGCAG	838						
OY	880	AAGGCGCAACTGCTTTCAAAGCACTGGCCCCCGGAGCAGAGAGAGAAATTGCACGCCGTGCGG	939						
Db	839	GAGGGGCGAGAGGGGCGAGGAGCAGAGAGAGAGGGGCGAAGAGCAGGAGGAGGGGCGAGAGGGG	898						
OY	940	GGCCTTCATCCGCCAGTACGAGACCACCTTGGAGGGCTCCTCAAAGCTGTGTGAGTCCGCC	999						
Db	899	CAGGACGAGAGAGAGGGGCGAGGAGGGGCGAGAGGGGCGAGGACAGGAGGAGAGGGGCGACGAG	958						
OY	1000	ATCCAGTCCATGAGAGAGGCCCGCAGATGCGCTCTTACTCCAGCAGCGCAAGAAGCACTGATC	1059						
Db	959	CAGGAGGAGGGGCGAGGAGGGGCGAGGAGCAGGAGAGGGGCGAGGAGGGGCGAG	1018						
OY	1060	AACCAAGTGGGGGCAATGTCTGAAGGTGAGAGCTGGCAGAGCGGCCGAGACCAAGCTTATAG	1119						
Db	1019	GAGCGAGGAGGAGGGGCGAGGAGCAGGAGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGCAG	1078						
OY	1120	AGCATGAGCAATTCTGTGTGAGCGTGGAGCACAGTCGCGCCAATGTTTCGAAMCAATCGAC	1179						
Db	1079	GGGCGAGGACAGAGAGAGAGGGGCGAGGAGGGGCGAGGAGGGCGAGAGCGAGGCGACAGAG	1138						
OY	1180	TTCCAAGCCGGGCGCCGCTGGGGATGTAAGAGGATGACACATGCGTTTGGATGGGAGAGAG	1239						
Db	1139	CAGGAGGGGCGAGGAGCAGGAGGGGCGAGGAGCAGGAGGAGGGGCGAGGAGGGGCGAGGAGG	1198						
OY	1240	GGCAATGCGGGGCTGGAGAGCGAGCGGCTGAGGTGCGACAG	1280						
Db	1199	GGGCGAGGAGGGCGAGGAGCAGGAGGGGCGAGGAGGAGGAGGGGCGAGG	1239						

RESULT 8  
US-09-620-925-1

```

;
; GENERAL INFORMATION:
;
; APPLICANT: Zuckermann et al.
;
; TITLE OF INVENTION: Compositions and Method
;                    Polynucleotide Delivery
;
; NUMBER OF SEQUENCES:

```

NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: California  
COUNTRY: U.S.A.

ZIP: 94608-2916  
COMPUTER READABLE FORM

```

1 MEDIUM TYPE: Floppy disk
2 COMPUTER: IBM PC compatible
3 OPERATING SYSTEM: PC-DOS/MS-DOS
4 SOFTWARE: Patentn Release #1.0, Version #1.30
5
6 CURRENT APPLICATION DATA:
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
10
```

1 FILING DATE: <Unknown>  
2 ATTORNEY/AGENT INFORMATION:  
3 NAME: Fujita, Sharon M., 459  
4 REGISTRATION NUMBER: 38,459  
5 REFERENCE/DOCKET NUMBER: 1218.000  
6 TELECOMMUNICATION INFORMATION:  
7 TELEPHONE: (510) 923-2706  
8 TELEFAX: (510) 655-3542  
9 INFORMATION FOR SEQ ID NO: 1:

```

?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 9600 base pairs
?      TYPE: nucleic acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
?      MOLECULE TYPE: DNA (genomic)
?      SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-620-925-1

Query Match          4.7%;   Score 66.6;   DB 4;   length 9600;
Best Local Similarity 45.5%;   Pred. No. 1.3e-06;
Matches 237;   Conservative 0;   Mismatches 284;   Indels 0;   Gaps 0;
```

QY	760	GTCTACGCGCATATCAACCCAGATGAGAGGAGTGTCCACGACCATTTGAGAGCAACAGCCGC	819
Db	719	GGGCGAGGACGAGGAGGGCGCAGAGCAGCAGAGAGAGGGGCGAGAGCAGAGAGAGGGGCGAGAG	778
QY	820	AGCAGAGACCACTGTTAAACCGAGAGTTCGAGACCTGTGTGCGCGTTTGGAGAGAGCC	879
Db	779	GGGCGAGAGAGGGGCGAGAGGGGCGAGAGCGAGAGAGAGGGGCGAGAGCAGAGAGAGGGCGAG	838
QY	880	AAGGGGCAACTGTTTCAACCACTGGCGCCCGAGGACGAGAGAGAAATTGCACGCGTGTGG	939
Db	839	GAGGGCGAGAGGGGCGAGAGCAGAGAGAGAGAGGGGCGAGAGCAGAGAGAGAGGGCGAGAGGGG	898
QY	940	GGCCTCATCGCCAGTAGCGAGACCACTTGGAGGCTCCTCAAGCTGTGTGATCCGCC	999
Db	899	CAGGAGCAGAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGCAGAGAGAGAGGGCGAGAG	958
QY	1000	ATCCAGTCCATGAGAGAGGCGCGCAGATGCGTCTCTACTCCACGACGAGCAAGAGCTGATC	1059
Db	959	CAGGAGAGGGGCGAGAGGGGCGAGAGCAGAGAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAG	1018
QY	1060	AACAGGTCGGGCAATGTTCGAAAGGTGAGACTGGCAGAGCGGCCGAGCCAGAGCTTATGAG	1119
Db	1019	GAGCAGAGAGAGGGGCGAGAGCGAGGAGGCGGCGAGAGAGGGGCGAGAGGCGAGCAGAGAG	1078
QY	1120	AGCATGAGAGCAATTCTCTGTGAGCCGTGAGACACGTGTGCCGAATTTGCAACCATTCGAC	1179
Db	1079	GGGCGAGAGCAGAGAGAGGGGCGAGAGGGGCGAGAGAGGGGCGAGAGACAGAGAGGGGCGAGAG	1138
QY	1180	TTTCAGCGCGGCCCTCTGGGGAATGGAAGAGATGACAGCATGGCTTTGGATGGGAGAG	1239
Db	1139	CAGAGGGGCGAGAGCAGAGAGGGGCGAGAGCAGAGAGAGAGGGGCGAGAGAGGGGCGAGAG	1198
QY	1240	GGCATATCGGGGCTGAGAGAGAGAGGGGCTGTGACGTGCCAGA	1280
Db	1199	GGGCGAGAGGGGCGAGAGCGAGAGGGGCGAGAGAGGGGCGAGAG	1239

RESULT 9  
US-07-884-811-15  
; Sequence 15, Application US/07884811

Patent NO. 5318921  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.  
TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
Addresser: Genotech, Inc

ADDRESS: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA

```

;
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/884,811
; FILING DATE: 19920518
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 755.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/952-9881
; TELEFAX: 415/225-3216
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-07-884-811-15

Query Match 4.7%; Score 66.6; DB 1; Length 10596;
Best Local Similarity 45.5%; Pred. No. 1.3e-06;
Matches 237; Conservative 0; Mismatches 284; Indels 0; Gaps 0;

;
; 760 GTGCAGGACAGTATACCCAGATGAGAGGAGGTCGTCAGACCATTTGAGAGCAACAGCCGC 819
; 2253 GGGCAGAGAGAGAGAGGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2312
; 820 AGACAGAGCAACTGTTAAACAGAGGTTCCGAGACCTGTGCGCGTTTGGAGAGAGCG 879
; 2313 GGGCAGAGAGGGGCGAGAGGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2372
; 880 AAGGCGCAACTGCTTCAAGCACTGCGCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 939
; 2373 GAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2432
; 940 GGCCCTATCCCGCACTGAGAGACCACTTGGAGGGCTCCTCAAGCTGTGTGAGTCCGCC 999
; 2433 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2492
; 1000 ATCCAGTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1059
; 2493 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2552
; 1060 AACAGAGTCCGGGCAATGTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1119
; 2553 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2612
; 1120 AGCATGAGCAATTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1179
; 2613 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2672
; 1180 TTCACAGCCGGGCGCGCTGGAGATGAAGAGATGAAGAGATGAAGATGAAGATGAAGAT 1239
; 2673 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2732
; 1240 GGCATGCGGGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1280
; 2733 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2773
;
; RESULT 10
; US-07-885-971-15
; Sequence 15, Application US/07885971
```

```

; Patent No. 5328837
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/885,971
; FILING DATE: 19920518
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-07-885-971-15

Query Match 4.7%; Score 66.6; DB 1; Length 10596;
Best Local Similarity 45.5%; Pred. No. 1.3e-06;
Matches 237; Conservative 0; Mismatches 284; Indels 0; Gaps 0;

;
; 760 GTGCAGGACAGTATACCCAGATGAGAGGAGTGTCCAGACCATTTGAGAGCAACAGCCGC 819
; 2253 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2312
; 820 AGACAGAGCAACTGTTAAACAGAGGTTCCGAGACCTGTGCGCGTTTGGAGAGAGCGC 879
; 2313 GGGCAGAGAGGGGCGAGAGGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2372
; 880 AAGGCGCAACTGCTTCAAGCACTGCGCCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 939
; 2373 GAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2432
; 940 GGCCCTATCCCGCACTAGAGAGACCACTTGGAGGGCTCCTCAAGCTGTGTGAGTCCGCC 999
; 2433 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2492
; 1000 ATCCAGTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1059
; 2493 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2552
; 1060 AACAGAGTCCGGGCAATGTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1119
; 2553 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2612
; 1120 AGCATGAGCAATTTCTGTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1179
; 2613 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2672
; 1180 TTCACAGCCGGGCGCGCTGGAGATGAAGAGATGAAGAGATGAAGATGAAGATGAAGATGA 1239
```





QY	820	AGACGAGGACCACTGTTTAAACCAGAGGTTCCGAACCCCTGTGCGGGGCTTTTGGAGAGCGC	879
Db	2313	GGGCAGGAGGGGCGAGGAGGGGCGCAGGACGACGAGGAGAGGGCGCAGGACGAGGAGGGGCGAG	2373
QY	880	AAGGGCGAACTGCTTCAMGCACTGCGCCCGGGAGCAGGAGGAGAATTGCGACGCGGTGCGG	939
Db	2373	GAGGGCGAGGAGGGGCGAGGACGAGGAGGAGGAGGGCGAGGACGAGGAGGAGGGGCGAGGAGGGG	2432
QY	940	GGCCTCATCCGCGCACTACGGAACCCACTTGGAGGGCTCCTCAAAAGCTGTGTGAAGTCCGCC	999
Db	2433	CAGGAGCAGGAGAGAGGGGCGAGGAGGGGCGAGGAGGCGAGGAGGAGGAGGAGGCGCAGAGAG	2492
QY	1000	ATCCAGTCATGAGGAGGCGCGCAGATGGCTCTTACCTCAGCAGGCGAAAGAGCACTGATC	1059
Db	2493	CAGGAGGAGGGGCGAGGAGGGGCGCAGGACGAGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAG	2555
QY	1060	AACCAAGTCGGGGCAATTGTCCAAAGGTGAGCTGGCAGAGCGCCCGGAGCCAGGCTTATGAG	1119
Db	2553	GAGCGAGGAGGAGGGGCGAGGACGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGCAGAG	2612
QY	1120	AGCATGAGGCAATTCTCTGTGAGCGTGAAGGAGCACTGGGCGCAATGTTGCCAACCATGAC	1179
Db	2613	GGGCAGGACACGAGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGACGAGGAGGGCGAGGAG	2672
QY	1180	TTCCAGCGGGGCGCGCTGGGGGTGAAGAGGATGACACATGCTTTTGGATGGGAGAG	1239
Db	2673	CAGGAGGGGCGAGGACGAGGAGGGGCGAGGACGAGGAGGGGCGAGGAGGGGCGAGGACGAGAG	2732
QY	1240	GGCATTCGGGGGCTGGAGGAGGAGCGCGCTGTGAGCTGCGCA	1280
Db	2733	GCGCAGGAGGGGCGAGGACGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAG	2773

RESULT 13  
 US-08-194-087-15  
 : Sequence 15, Application US/08194087  
 : Patent No. 5879910  
 : GENERAL INFORMATION:  
 : APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.  
 : TITLE OF INVENTION: 'HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS'  
 : NUMBER OF SEQUENCES: 21  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Genentech, Inc.  
 : STREET: 460 Point San Bruno Blvd  
 : CITY: South San Francisco  
 : STATE: California  
 : COUNTRY: USA  
 : ZIP: 94080  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: patin (Genentech)  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/194,087  
 : FILING DATE: 18-MAY-1992  
 : CLASSIFICATION: 435  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER:  
 : FILING DATE:  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Dreger, Ginger R.  
 : REGISTRATION NUMBER: 33,055  
 : REFERENCE/DOCKET NUMBER: 779  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 415/225-3216  
 : TELEFAX: 415/952-9881  
 : TELE: 910/371-7168  
 : INFORMATION FOR SEQ ID NO: 15:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 10596 bases  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single

[illegible]

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 755,779P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEFAX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: Single
; TOPOLOGY: linear
; PCT-US93-04648-15

Query Match 4.7%; Score 66.6; DB 5; Length 10596;
Best Local Similarity 45.5%; Pred. No. 1.3e-06;
Matches 237; Conservative 0; Mismatches 284; Indels 0; Gaps 0;

QY 760 GTGCAGGCGATGATCCAGATGAGGAGGCTGCGCAGACCAATTGAGGCAACAGCCGC 819
   |||||
Db 223 GGGCAGAGAGAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2312
   |||||
QY 820 AGACAGAAACCACTGTTAAACAGAGGTTTCGAGACCCCTGTGCGCGTTTTCAGAGAGCGC 879
   |||||
Db 2313 GGGCAGAGAGAGAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2372
   |||||
QY 880 AAGGCGGAATGCTTTAAAGCAGCTGCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 939
   |||||
Db 2373 GAGGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2432
   |||||
QY 940 GGCCTCATCCGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 999
   |||||
Db 2433 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2492
   |||||
QY 1000 ATCCAGTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1059
   |||||
Db 2493 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2552
   |||||
QY 1060 AACAGAGTGGGGGCAATGTCGAAGGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1119
   |||||
Db 2553 GAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2612
   |||||
QY 1120 AGCATGAGCAATTTCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1179
   |||||
Db 2613 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2672
   |||||
QY 1180 TTCAGAGCCGGGCGCCCTGTGGGATGAAGAGAGATGACGACATGCTTTGGATGGGAGAG 1239
   |||||
Db 2673 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2732
   |||||
QY 1240 GGCATGCGGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1280
   |||||
Db 2733 GGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2773
   |||||

RESULT 15
US-09-249-585A-4/c
; Sequence 4, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/00905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus

```

```

; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)...(1926)
; OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4

Query Match 4.5%; Score 64; DB 4; Length 1926;
Best Local Similarity 44.9%; Pred. No. 3.2e-06;
Matches 284; Conservative 0; Mismatches 345; Indels 3; Gaps 1;

QY 653 TCGGCGCCACAGAGACTGTGAGTGCCCTCTGCGCCACCAATTTCACAAACGCCAGAGA 712
   |||||
Db 1179 TCTGGCCCTACTACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1120
   |||||
QY 713 GTGAGCTGAGCATGGATCGCATCGCATGCTGTGCGGGCAATGACCTGTGACAGCATGA 772
   |||||
Db 1119 TGGAGAGGGGACCCAGAAAGTGTGAAGGGGGGACCCAGAAAGTGCAGAGATGGGGC 1060
   |||||
QY 773 TCACCCAGATGAGAGAGAGTGTGCGACACATTGAGAGACACAGCCGACAGAGCAAC 832
   |||||
Db 1059 CGCCGAGGTGATGAGAGAGAGTGTGCGCCGAGAGGTGATGAGAGAGTGTGCGGCGCCGAGGTGA 1000
   |||||
QY 833 TGTTAACAGAGGTTTCGAGACCCCTGTGCGCGTTTGGAGAGAGCCCAAGGGCACTGC 892
   |||||
Db 999 CGAGAGAGCTGGGGCGCGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 940
   |||||
QY 893 TTCAAGCACTGCGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 952
   |||||
Db 939 ACGAGAGCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 880
   |||||
QY 953 AGTAGAGAGACCACTTTGAGAGGCTCTCAAGCTGTGTGAGTCCGCTATCCATGATG 1012
   |||||
Db 879 AGGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 820
   |||||
QY 1013 AGGAGCGGAGATGGCTCTCTACCTCAGAGGCAAGAGAGTGTATCAACAGATCGGGG 1072
   |||||
Db 819 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 760
   |||||
QY 1073 CAATGTCGAAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1132
   |||||
Db 759 ACGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 700
   |||||
QY 1133 TCTCTGTGAGCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1192
   |||||
Db 699 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 643
   |||||
QY 1193 CGGCTGGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1252
   |||||
Db 642 AGGAGCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 583
   |||||
QY 1253 TGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1284
   |||||
Db 582 AGGAGCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 551
   |||||

RESULT 16
US-09-130-114-2/c
; Sequence 2, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; FILE REFERENCE: 0867/1D903JUS1
; CURRENT APPLICATION NUMBER: US/09/130,114
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1931
; TYPE: DNA

```

ORGANISM: EBNA  
US-09-130-114-2

Query Match 4.5%; Score 64; DB 2; Length 1931;

Best Local Similarity 44.9%; Pred. No. 3.2e-06;  
Matches 284; Conservative 0; Mismatches 345; Indels 3; Gaps 1;

```

QY 653 TCGGCGCCCAAGAGACTGTGAGTGCGCCCTCTGCCCACATTATCAAGCCGCAAGA 712
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1179 TCTGGGCTCTACTACTGATGATGATGATGATGATGATGATGATGATGATGATG 1120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 713 GTGAGCTGAGCGATGCGATGCGATGCTGTGCGCGGCAATGACCTGTGTGAGCAGTGA 772
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1119 TGGAGAGGGGGACCGGAAAGTGTCTAAGGGGGGACCGGAAAGTGTGAGAGATGGGCG 1060
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 773 TCACCCAGATGAGAGAGTGTGTCAGACCTTGTGAGGACAAACAGCCGCAAGAGCAAC 832
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1059 CGCCGAGAGTGTATGAGAGAGCTGGGGCCGAGGTGATGAGAGAGCTGGGGCGGAGTGA 1000
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 833 TGTAAACCAAGAGTTCAGAGACCTGTGCGCGGTTTGTGAGAGACCGCAAGGCAACTGC 892
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 999 CGGAGAGAGCTGGGGCCGAGGTGAGAGAGAGAGACGGGAGAGACGAGGAGAGAG 940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 893 TTCAAGCACTGCGCCGGGAGACAGAGAGAGAAATTGCAAGCGCGGTGCGGCGCTCATCCGCC 952
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 939 ACGAGAGCGGGAGAGACGGGGAGAGACGGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 880
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 953 AGTAGAGAGACCACTTGTGAGGGCTCTCAAAAGCTGTGAGAGTCCGCCATCCAGTCCATAG 1012
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 879 AGGAGCGGGAGAGACAGAGACGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1013 AGGAGCCCAAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1072
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 819 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 760
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1073 CAATGTCCAAAGTGGAGCTGCGAGACAGCGCCGAGACCGAGCTGTGAGAGCATGAGCAAT 1132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 759 ACGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1133 TCTCTGAGCGTGTGAGAGACAGCGCCGAATGTTGCAACCATTCAGTCCAGCGGGCGG 1192
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 699 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1193 CCGCTGGGATGAAAGAGATGACAGCATGCTTTGGATGGGAGAGAGCGCAATGCGGGCG 1252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 642 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1253 TGGAGAGAGAGCGCTGTGAGAGTGTCCAGAAAGCC 1284
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 582 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 551
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

# RESULT 17

US-08-232-463-14/C  
Sequence 14, Application US/08232463  
Patent No. 5670367

## GENERAL INFORMATION:

APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOXLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA

## COUNTRY: USA

ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232.463

FILING DATE:  
CLASSIFICATION: 435

Prior APPLICATION DATA:  
APPLICATION NUMBER: US/07/935.313

FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6

FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

CLONE: pTZgpt-F15

US-08-232-463-14

Query Match 4.2%; Score 59.8; DB 1; Length 7218;  
Best Local Similarity 6.0%; Pred. No. 4.7e-05;

Matches 25; Conservative 224; Mismatches 166; Indels 0; Gaps 0;

```

QY 981 AAGCTGTGTGAGTCCGCCATCCAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1040
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1455 AGAGATAGAGAAATTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1041 GCAGCAAGAGAGCTGTCAACAGTCCGGGCAATGTCGAAGTGGAGTGGCGAGAGAGC 1100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1395 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1101 GCCGAGCAGAGCTATGAGAGATGAGCAATCTCTGTGAGCGTGAGAGAGAGAGAGAG 1160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1335 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1161 AATGTCCAGACCATGACTTCCAGCGCGCGCGCTGGGATGAGAGATGAGAGAT 1220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1275 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1221 GCGTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1215 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1281 AGGCTCAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1155 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1096
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1341 GGGATGCTAGAGATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1095 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

# RESULT 18

US-08-728-323A-1  
Sequence 1, Application US/08728323A  
Patent No. 5948676

## GENERAL INFORMATION:

APPLICANT: Chang, Yuan  
APPLICANT: Bohenzky, Roy A.  
APPLICANT: Russo, James J.  
APPLICANT: Edelman, Isidore S.  
APPLICANT: Moore, Patrick S.  
TITLE OF INVENTION: Immediate Early Protein From Kaposi's  
TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA

```

; TITLE OF INVENTION: Encoding Same And Uses Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/728,323A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; RESISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/52268/JPM/NSC/SKS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3489 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3489
; US-08-728-323A-1

Query Match
Best Local Similarity 3.8%; Score 54.2; DB 2; Length 3489;
Matches 257; Conservative 0; Mismatches 338; Indels 0; Gaps 0;

QY 701 AACGCCAAGAAGTGTGAGCTGAGCGATGCGGATCGGCTGGGGCGCATGACCGTG 760
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1976 AGCAGCAGCAGATGAGCAGCAGCAGCAGATGAGCAGCAGCAGATGAGCAGCAGATG 2035
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 761 TGCAGCAGTGTATCACCCAGATGAGAGAGGTGTGCCAGACCATTTAGAGACAACGCCGA 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2036 ACACAGCAGCAGATGATGACAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGCAGATG 2095
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 821 GACAGAGAAGCTGTAAACCAAGATTCAGAGACCTGTGCGCGCTTTGGAGAGCGCA 880
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2096 AGCAGAGCAGCAGATGAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGATGAGCAGC 2155
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 881 AGGGCGAAGCTGTTCAAGCACTGCCCCGGAGCAGAGAGAGAGTTGCAGCGCGTGGCGG 940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2156 AGCAGAGATGAGCAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGATGAGCAGCAGCAG 2215
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 941 GCCTATCCGCCAGTATGAGAGACCACTTGGAGGGCTCTCTAAAGCTGTGAGATCCGCCA 1000
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2216 ATGACAGCAGCAGAGATGAGCAGCAGCAGCAGATGAAACAGAGCAGCAGAGAGAGC 2275
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1001 TCCAGTCCATGAGAGAGCGCGAGATGGCTCTCACTCCAGCAGCGCAAGAGAGCTGATCA 1060
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2276 AGGAGCAGCAGAGAGAGCAGCAGCAGAGATTAGAGAGCAGCAGAGATTAGAGAGATC 2335
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1061 ACAAGTCGGGGCAATGTGGAAGTGTGAGCTGCGAGACCGCCAGCGCTATGAGA 1120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2336 AGGAGCAGAGAGATTAGAGAGCAGCAGCAGAGATTAGAGAGCAGCAGAGATTAGAG 2395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1121 GCATGAGCAATTTCTCTGTGAGCGTGTGAGCAGCTGGCCCAATGTTTGCAGAACATGACT 1180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2396 AGCAGAGCAGAGAGATTAGAGAGCAGCAGCAGAGATTAGAGAGCAGAGAGATTAG 2455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1241 GCATGCGGGGCTGTGAGAGAGAGCGGCTGTGACGTGCAGAAAGCTTACGCTGCA 1295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 1181 TCCAGCCGGGGCGCGCTGGCGATGAAGAGATGACAGCATGCTTTGCATGGGAGAG 1240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2456 AGGAGCAGAGCAGCAGAGATTAGAGAGCAGCAGAGATTAGAGAGCAGCAGAGAGT 2515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1241 GCATGCGGGGCTGTGAGAGAGAGCGGCTGTGAGCTGCCAGAGAGCTCAGCGCTGCA 1295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2516 TAGAGAGCAGCAGAGCTGGAAGAGCAGCAGAGAGCTGGAAGAGCAGCAGCAGCA 2570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 19
US-09-298-568-1
; Sequence 1, Application US/09298568
; Patent No. 6322792
; GENERAL INFORMATION:
; APPLICANT: Kieff, Elliott D.
; APPLICANT: Ballestas, Mary E.
; APPLICANT: Kaye, Kenneth M.
; TITLE OF INVENTION: RHADINO VIRUS DNA ACTS IN TRANS ON A UNIT OF RHADINO
; FILE REFERENCE: 16412-10001R
; CURRENT APPLICATION NUMBER: US/09/298,568
; EARLIER FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,422
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 1
; LENGTH: 3489
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-298-568-1

```

```

Query Match
Best Local Similarity 3.8%; Score 54.2; DB 4; Length 3489;
Matches 257; Conservative 0; Mismatches 338; Indels 0; Gaps 0;

QY 701 AACGCCAAGAAGTGTGAGCTGAGCGATGCGGATCGGCTGGGGCGCATGACCGTG 760
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1976 AGCAGCAGCAGATGAGCAGCAGCAGCAGATGAGCAGCAGCAGATGAGCAGCAGATG 2035
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 761 TGCAGCAGTGTATCACCCAGATGAGAGAGGTGTGCCAGACCATTTAGAGACAACGCCGA 820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2036 ACACAGCAGCAGATGATGACAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGCAGATG 2095
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 821 GACAGAGAAGCTGTAAACCAAGATTCAGAGACCTGTGCGCGCTTTGGAGAGCGCA 880
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2096 ACACAGCAGCAGATGATGAGCAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGCAGATG 2155
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 881 AGGGCGAAGCTGTTCAAGCACTGCCCCGGAGCAGAGAGAGAGTTGCAGCGCGTGGCGG 940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2156 AGCAGAGATGAGCAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGATGAGCAGCAGCAG 2215
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 941 GCCTATCCGCCAGTATGAGAGACCACTTGGAGGGCTCTCTAAAGCTGTGAGATCCGCCA 1000
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2216 ATGACAGCAGCAGAGATGAGCAGCAGCAGCAGATGAAACAGAGCAGCAGAGAGAGC 2275
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1001 TCCAGTCCATGAGAGAGCGCGAGATGGCTCTCACTCCAGCAGCGCAAGAGAGCTGATCA 1060
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2276 AGGAGCAGCAGAGAGAGCAGCAGCAGAGATTAGAGAGCAGCAGAGATTAGAGAGATC 2335
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1061 ACAAGTCGGGGCAATGTGGAAGTGTGAGCTGCGAGACCGCCAGCGCTATGAGA 1120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2336 AGGAGCAGAGAGATTAGAGAGCAGCAGCAGAGATTAGAGAGCAGCAGAGATTAGAG 2395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1121 GCATGAGCAATTTCTCTGTGAGCGTGTGAGCAGCTGGCCCAATGTTTGCAGAACATGACT 1180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2396 AGCAGAGCAGAGAGATTAGAGAGCAGCAGCAGAGATTAGAGAGCAGAGAGATTAG 2455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1181 TCCAGCCGGGGCGCGCTGGGAGTGAAGAGATGAGCAATGCTTTGGATGGGAGAG 1240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2456 AGGAGCAGAGAGAGATTAGAGAGCAGCAGCAGAGATTAGAGAGCAGCAGAGAGT 2515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1241 GCATGCGGGGCTGTGAGAGAGAGCGGCTGTGACGTGCAGAAAGCTTACGCTGCA 1295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 20  
US-08-770-379-20/C  
Sequence 20, Application US/08770379  
Patent No. 5849564  
GENERAL INFORMATION:  
APPLICANT: Chang, Yuan  
APPLICANT: Bohenzky, Roy A.  
APPLICANT: Russo, James J.  
APPLICANT: Edelman, Isidore S.  
APPLICANT: Moore, Patrick S.  
TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED  
HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/770,379  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 52342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEO ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32207 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-770-379-20

Query Match 3.88; Score 54.2; DB 2; Length 32207;  
Best Local Similarity 43.28; Pred. No. 0.0015;  
Matches 257; Conservative 0; Mismatches 338; Indels 0; Gaps 0;

DB 2516 TAGAGGAGCAGAGGTGTGAAGACCAAGACGAGGAGTGTGAAGACCAAGACGAGGA 2570  
DB 701 AACGCCAGAAAGTGTGAGCTGAGCGATGCGATGCGATGCTGTGCGGGCAATGACCGTG 760  
DB 20021 AGCAGCAGCAGAGATGAGCAGCAGCAGATGAGCAGCAGCAGATGAGCAGCAGATG 19962  
DB 761 TGCAGCAGATGATCAACCATGAGAGAGGTGTGCGCAGACATTTGAGACACACGCCGA 820  
DB 19961 AGCAGCAGCAGATGAGCAGCAGCAGATGAGCAGCAGCAGATGAGCAGCAGCAGATG 19902  
DB 821 GACAGAGCAGACTGTTAAACCGAGAGGTTCGAGACCTGTGCGCGTTTGTGAGAGCGCA 880  
DB 19901 AGCAGAGCAGCAGATGAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGATGAGCAGC 19842  
DB 881 AGGCGAAGCTGCTTCAAGCACTGGCCCGGAGCAGAGAGAGAAATGTCACCGCGTGGG 940  
DB 19841 AGCAGAGATGAGCAGCAGCAGATGAGCAGCAGCAGATGAGCAGCAGCAGCAGCAGG 19782  
DB 941 GCTTCATCCGCGCAGTACGAGACCACTTGGAGGGCTCTCAAAAGCTGTGTGAGTCCGCCA 1000  
DB 19781 ATGAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGCAGCAG 19722

QY 1001 TCCAGTCCATGAGAGAGCGCGCAGATGCGCTCTACTCCAGCAGCAAGAGCTGATCA 1060  
DB 19721 AGGAGCAGCAGAGAGAGCAGCAGAGCAGAGATTAGAGAGCAGAGCAGAGATTAGAGATC 19662  
QY 1061 ACAAGTTCGGGGCAATGTCGAAGGTGAGCTGGCAGAGCGCCGGAGCCAGCTATGAGA 1120  
DB 19661 AGGAGCAGAGATTAGAGAGCAGCAGAGAGATTAGAGAGCAGCAGAGATTAGAGAG 19602  
QY 1121 GCATGAGCAATTCTCTGTGAGCCGTGAGCAGCAGTGGCGCAATGTTGCCAACCATTGACT 1180  
DB 19601 AGCAGAGCAGAGAGATTAGAGAGCAGAGCAGAGATTAGAGAGCAGAGCAGAGATTAG 19542  
QY 1181 TCCAGCGCGCGCGCTGGGATGAAGAGATGAGCAGCAGTGGCTTTGATGGGAGAGG 1240  
DB 19541 AGGAGCAGAGCAGAGATTAGAGAGCAGCAGAGCAGAGATTAGAGAGCAGAGCAGAG 19482  
QY 1241 GCAATGCGGGGCTGAGAGAGAGCGGCTGGAGCTGCCAGAGGCTCAGGCTTCA 1295  
DB 19481 TAGAGAGCAGAGAGGTGAAGAGCAAGAGCAGAGAGGTGAAGAGCAAGAGCAGGA 19427

Search completed: December 3, 2002, 09:13:33  
Job time : 194 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

December 3, 2002, 09:05:22 ; Search time 64 Seconds  
(without alignments)

8610.941 Million cell updates/sec

Title: US-09-908-988B-1

Perfect score: 1431  
1 aagagagctgtgacagagtggtt.....ataaagactcaagtgctcc 1431

Scoring table: IDENTITY\_NUC  
Gap 10.0, Gapext 1.0

Searched: 341543 seqs, 192557720 residues

Total number of hits satisfying chosen parameters: 683086

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCR\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCRUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1431	100.0	1431	US-09-908-988B-1	Sequence 1, Appl1
2	484.6	33.9	1597	US-09-908-988B-5	Sequence 5, Appl1
3	478.8	33.5	1781	US-09-764-864-21	Sequence 21, Appl1
4	450.8	31.5	2580	US-09-908-988B-3	Sequence 3, Appl1
5	439.6	30.7	1762	US-09-764-864-34	Sequence 34, Appl1
6	229.2	16.0	587	US-09-764-864-493	Sequence 493, Appl1
7	162	11.3	446	US-09-764-864-482	Sequence 482, Appl1
8	141.6	9.9	1039	US-09-880-192-25	Sequence 25, Appl1
9	59.6	4.2	1852	US-09-969-852-4	Sequence 4, Appl1
10	50.8	3.5	1400	US-09-764-864-56	Sequence 56, Appl1
11	50.4	3.5	3441	US-09-880-192-41	Sequence 41, Appl1
12	48	3.4	7065	US-09-981-496-115	Sequence 115, Appl1
13	48	3.4	7065	US-09-874-923-115	Sequence 115, Appl1
14	47.4	3.3	768	US-09-938-842A-812	Sequence 812, Appl1
15	46.4	3.2	280	US-09-923-876-5584	Sequence 5584, Appl1
16	45.2	3.2	1969	US-09-864-761-3972	Sequence 3972, Appl1
17	43.8	3.1	1464	US-09-771-161A-54	Sequence 54, Appl1
18	43.8	3.1	2518	US-09-771-161A-53	Sequence 53, Appl1
19	43.8	3.1	659158	US-09-771-208-20	Sequence 20, Appl1

20	43.6	3.0	319	US-09-960-352-6282	Sequence 6282, Appl1
21	43.6	3.0	1044	US-09-969-708-167	Sequence 167, Appl1
22	43.6	3.0	1044	US-09-880-107-2241	Sequence 2241, Appl1
23	43.6	3.0	4854	US-09-963-875-2	Sequence 2, Appl1
24	43.4	3.0	2108	US-09-962-832-225	Sequence 225, Appl1
25	42.8	3.0	765	US-09-759-143-524	Sequence 524, Appl1
26	42.8	3.0	765	US-09-780-669-524	Sequence 524, Appl1
27	42.8	3.0	765	US-09-822-827-524	Sequence 177, Appl1
28	42.8	3.0	1119	US-09-759-143-177	Sequence 177, Appl1
29	42.8	3.0	1119	US-09-780-669-177	Sequence 177, Appl1
30	42.8	3.0	1119	US-09-030-606-177	Sequence 177, Appl1
31	42.8	3.0	1119	US-09-822-827-177	Sequence 177, Appl1
32	42.8	3.0	1119	US-09-115-453-177	Sequence 177, Appl1
33	42.8	3.0	1214	US-09-759-143-225	Sequence 225, Appl1
34	42.8	3.0	1214	US-09-780-669-225	Sequence 225, Appl1
35	42.8	3.0	1214	US-09-822-827-225	Sequence 225, Appl1
36	42.8	3.0	1214	US-09-115-453-225	Sequence 225, Appl1
37	42.8	3.0	1215	US-09-759-143-326	Sequence 326, Appl1
38	42.8	3.0	1215	US-09-780-669-326	Sequence 326, Appl1
39	42.8	3.0	1215	US-09-822-827-326	Sequence 326, Appl1
40	42.8	3.0	1265	US-09-759-143-173	Sequence 173, Appl1
41	42.8	3.0	1265	US-09-780-669-173	Sequence 173, Appl1
42	42.8	3.0	1265	US-09-030-606-173	Sequence 173, Appl1
43	42.8	3.0	1265	US-09-822-827-173	Sequence 173, Appl1
44	42.8	3.0	1265	US-09-115-453-173	Sequence 173, Appl1
45	42.8	3.0	1350	US-09-759-143-616	Sequence 616, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-908-988B-1  
Sequence 1, Application US/09908988B  
Patent No. US20020127690A1  
GENERAL INFORMATION:  
APPLICANT: OLSON, JEFFREY A.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES  
FILE REFERENCE: MYOG:02808  
CURRENT APPLICATION NUMBER: US/09/908, 988B  
CURRENT FILING DATE: 2000-07-18  
PRIOR APPLICATION NUMBER: 60/219, 020  
PRIOR FILING DATE: 2000-07-18  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 1431  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (199)..(1296)  
US-09-908-988B-1  
Query Match  
Best Local Similarity 100.0%; Score 1431; DB 10; Length 1431;  
Matches 1431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAGAGCTGACAGAGCTGCTGGAATACACAGGGGTGAGAGAGCTGTTAGGGAGG 60  
DB 1 AAGGAGCTGACAGAGCTGCTGGAATACACAGGGGTGAGAGAGCTGTTAGGGAGG 60  
QY 61 GACGAGCTCTTCACAGAGGACCAATAGCCGGATCCCAAGATCCAGTACGCTTAAC 120  
DB 61 GACGAGCTCTTCACAGAGGACCAATAGCCGGATCCCAAGATCCAGTACGCTTAAC 120  
QY 121 TGACCGAGGAGGAGGTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
DB 121 TGACCGAGGAGGAGGTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
QY 181 GCTCCAGAGGCGCCGGGAGTGAACCTTCACGCTGGGTTTCAAGCCGCTGCTGAGGAGTCCG 240

```

Db 181 GCTCCAGAGCGCGGGGATGAACTTACAGCTGGGTTTCAAGCGGTGCTAGGGGATGCG 240
QY 241 CACACATGAGAACACTTGGAGAGAGAGCTCATTTGGCCCATCTGCTGGAGATGTTCTCC 300
Db 241 CACAACATGAGAACACTTGGAGAGAGAGCTCATTTGGCCCATCTGCTGGAGATGTTCTCC 300
QY 301 AAGCCCGTGTGATCTGCTGCTGCGCAACACAACTGTCGCGCAAGTGTGCCAAGGAGCTC 360
Db 301 AAGCCCGTGTGATCTGCTGCTGCGCAACACAACTGTCGCGCGCAAGTGTGCCAAGGAGCTC 360
QY 361 TTCCAGGCTCTAATCTCTGTGGCAATCCCGGGCTCCAAACGGTGTCTTCAGAGAGA 420
Db 361 TTCCAGGCTCTAATCTCTGTGGCAATCCCGGGCTCCAAACGGTGTCTTCAGAGAGA 420
QY 421 CGTTCCGATGCGCCATCTTGTAGGAGAGGTTGCTCTGACAGAGCATGCTGTATGCG 480
Db 421 CGTTCCGATGCGCCATCTTGTAGGAGAGGTTGCTCTGACAGAGCATGCTGTATGCG 480
QY 481 CTGAGCGGAACCTGCTAGTGGAGAACATCTGACATCTACAGAGAGTCTCTCCG 540
Db 481 CTGAGCGGAACCTGCTAGTGGAGAACATCTGACATCTACAGAGAGTCTCTCCG 540
QY 541 CCATGCAACGCCAAGGCTGAGACAGACATCTATGTGAGAGAGCAGAGAGAGATC 600
Db 541 CCATGCAACGCCAAGGCTGAGACAGACATCTATGTGAGAGAGCAGAGAGAGATC 600
QY 601 AACATCTACTGCTGAGCTGGAGGAGGCTGCTCTCTCTGCAAGGTTTTCGGGCC 660
Db 601 AACATCTACTGCTGAGCTGGAGGAGGCTGCTCTCTCTGCAAGGTTTTCGGGCC 660
QY 661 CACAGAGACTGTAGAGTGGCCCTCTGCCCACATTTACAAACGCCAGAGAGTGAAGT 720
Db 661 CACAGAGACTGTAGAGTGGCCCTCTGCCCACATTTACAAACGCCAGAGAGTGAAGT 720
QY 721 AGCGATGGCATCGGATCTGTGGCCGGGCAATGACCTGTGAGAGCATGATCACCAG 780
Db 721 AGCGATGGCATCGGATCTGTGGCCGGGCAATGACCTGTGAGAGCATGATCACCAG 780
QY 781 ATGAGAGAGTGTGCTCAGACCATTTAGAGACAACAGCCGAGAGAGAGCAACTGTTAAAC 840
Db 781 ATGAGAGAGTGTGCTCAGACCATTTAGAGACAACAGCCGAGAGAGAGCAACTGTTAAAC 840
QY 841 CAGAGTTCAGAGACCTGTGCGGCTTTTGGAGAGCGCAAGGCGCAACTGCTTCAAGCA 900
Db 841 CAGAGTTCAGAGACCTGTGCGGCTTTTGGAGAGCGCAAGGCGCAACTGCTTCAAGCA 900
QY 901 CTGGCCCGGAGAGAGAGAGAGTGGAGCGGCTGCGGGGCTCATCCGCCAGTACGGA 960
Db 901 CTGGCCCGGAGAGAGAGAGAGTGGAGCGGCTGCGGGGCTCATCCGCCAGTACGGA 960
QY 961 GACCACTTGGAGGCTCTCTAAAGCTGTGAGTCCGCAATCAAGTCAATGAGAGAGCG 1020
Db 961 GACCACTTGGAGGCTCTCTAAAGCTGTGAGTCCGCAATCAAGTCAATGAGAGAGCG 1020
QY 1021 CAAATGGCTCTTCACTTCCAGAGCAAGCAAGCAAGCAAGTCAATGAGAGAGTCTG 1080
Db 1021 CAAATGGCTCTTCACTTCCAGAGCAAGCAAGCAAGTCAATGAGAGAGTCTG 1080
QY 1081 AAGGTGAGAGTGGAGAGAGCGCGAGAGCAAGTGTAGAGATGAGAGCAATTTCTGTG 1140
Db 1081 AAGGTGAGAGTGGAGAGAGCGCGAGAGCAAGTGTAGAGATGAGAGCAATTTCTGTG 1140
QY 1141 AGCGTGGAGAGCTGCGCGAATGTTGCGAACATCGACTTCCAGCCGCGCGCTGGG 1200
Db 1141 AGCGTGGAGAGCTGCGCGAATGTTGCGAACATCGACTTCCAGCCGCGCGCTGGG 1200
QY 1201 GATGAAGAGATGACGATGCTTTGGATGGGAGAGAGGCAATGCGGCGCTGAGAGAG 1260
Db 1201 GATGAAGAGATGACGATGCTTTGGATGGGAGAGAGGCAATGCGGCGCTGAGAGAG 1260
QY 1261 GAGCGGCTGAGAGTGCAGAGAGGCTGAGAGTGAACCGGATCTGATCAGAGCGC 1320

```

```

Db 1261 GAGCGGCTGAGAGTGCAGAGAGGCTCAGAGGCTGACATGACCCGACTGTATCCAGAGCGC 1320
QY 1321 ACACCCGAAGCGGAGAGCAAGAGGATGCTGAGAGATCTGCGAGAGACCAACCGGCCACCAA 1380
Db 1321 ACACCCGAAGCGGAGAGCAAGAGGATGCTGAGAGATCTGCGAGAGACCAACCGGCCACCAA 1380
QY 1381 GCTCGGCTTCCCGCCCCCGGGAAGGTTCTCAATTAAGGACTCAAGTGTCC 1431
Db 1381 GCTCGGCTTCCCGCCCCCGGGAAGGTTCTCAATTAAGGACTCAAGTGTCC 1431

RESULT 2
US-09-908-988b-5
; Sequence 5, Application US/09908988B
; Patent No. US20020127690A1
; GENERAL INFORMATION:
; APPLICANT: OLSON, ERIC
; APPLICANT: SPENCER, JEFFREY A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES
; FILE REFERENCE: MYOG:0280S
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/219,020
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1597
; TYPE: DNA
; ORGANISM: Mus musculus
; NAME/KEY: CDS
; LOCATION: (299)..(1327)
US-09-908-988b-5

Query Match 33.9%; Score 484.6; DB 10; Length 1597;
Best Local Similarity 69.2%; Pred. No. 2.6e-107;
Matches 677; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

QY 226 CTGCTAGGGAGTGCAGACATGAGACAACTTGGAGAGAGCTCATTTGCCCATCTGC 285
Db 317 CTGATTCCTGATGAGAAACCTATGAGAGAACTGAGAGAGAGAGTGTGCCATCTGC 376
QY 286 CTGGAGATGTTTCCAAAGCCGCTGATCTTCCCTGCCCAACAACCTGTGCGCAG 345
Db 377 CTGGAGATGTTTCCAAAGCCGCTGATCTTCCCTGCCCAACAACCTGTGCGCAG 436
QY 346 TGTGCCAAGAGCTCTTCCAGGCTCTTAATCTCTGTGGCAATCCCGGGGCTCCACAAG 405
Db 437 TGTGCCAAGAGATCTTCCAGGCTCTTAATCTCTGTGGCAATCCCGGGGCTCCACAAG 496
QY 406 GTGTTTCAAGAGAGAGTTCGAGTCCCATCTTGTAGGAGAGAGTGTGCTGAGAGAG 465
Db 497 TCCATGTCTGAGAGTGTCTTCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556
QY 466 CATGTGTCTATGAGGCTGAGGAGGAACTGTAGTGAAGAACATCATGTGATACAG 525
Db 557 CACGGGAGTGTACGCTGCTGAGAGAGAGCTGTGAGTGAAGAAATCATGTGATACAG 616
QY 526 CAGAGTCTCCCGGCACTGACGCGCAAGGCTGAACAGACACTCATGTGTAGAGAGAC 585
Db 617 CAGAGTCTCTCCAG--TCCGCCCCGTGACAGAAAGGAGCAGCCAGTGTGCAAGAGAC 673
QY 586 GAGAGCGAAGATTAACATCTACTGCTGAGCTGCGAGAGTCCCACTGCTCTCTGCG 645
Db 674 GAAAGCGAAGATTAACATCTACTGCTGAGCTGCGAGAGTCCCACTGCTCTCTGCG 733
QY 646 AAGGTTTTCGCGCCCAAGAGCTGTGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 705
Db 734 AAGGTTTTCGCGCCCAAGAGCTGTGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 793
QY 706 CAGAAAGTGAAGTGAAGGATGAGATCGGATGTGTTGGCGGGAATGACCGTGTGAG 765

```



Db	794	CAGAAAGACTGAGCTAGTACTGATCTGCACTCTCTCATCTGCTGTGGCGGGGAACGACCCAGTCCAG	853
QY	766	GCAGTAGTCAACCCAGATGAGGAGGAGGTGTGCCACAGACCATTGGAGGACAACAGCCCGACAGAC	825
Db	854	ACGATCATCTCTCACTGCTGGAGGAGCTGCTGACAGTAGTACCAAGAGAAATGCCACCAAGTGC	913
QY	826	AAGCAACTGTTTAACACAGAGTTCGAGAACCTCTGTGCGCGGTTTTTGGAGGAGCGCAAGGCG	885
Db	914	AAGGAGGAGCTGAGTACAGAAATTTTGACACCTCTTACGCCATCCTTGATGAGAAAGAAAGCG	973
QY	886	GACATGCTTCAAGCACTGGCCCGGGGAGAGGAGAGAAAGTTGACAGCCGTCGCGGGGGCTC	945
Db	974	GAGCTGCTGACGCGGATCACGCAAGAGACAGAGAGAAAGCTGGGGCTTCATTCGAGGCTCTG	1033
QY	946	ATCCGCCAGTACGAGAGACCACTTGGAGGAGGCGCTCTCAAAGCTGGTGGATGTCGCCATCCAG	1005
Db	1034	ATCCCTCCAGTACAGGAGAGCAGCTGGAAAAGTCCACCAGCTTGGAGACCGGCATCCAG	1093
QY	1006	TCCATGAGAGAGACCCGACATGTGCTCTCTACTCTCAGCAGGCAAAAGGAGCTGATTAACAAG	1055
Db	1094	TCCCTGGATAGACCCCGAGGGGCTACTCTTCTTCAAGTCCCAAGCGCTCATCAAGAGCG	1153
QY	1066	GTCGGGGCAATGTCGAAGGTGGAGCTGGGAGCGAGGAGGCGCGGAGCCAGCTATGAGAGCATG	1125
Db	1154	ATTGTAGAAAGCCTCCCAAGGGCTGCCAGCTGGGGAGACAGAGCAAGGCTTTGAGAAACATG	1213
QY	1126	GAGCAATTCCTCTGTGAGCGCTGGAGACGACGCTGGCGGAATGTTGGCAACATCGACTTCCAG	1185
Db	1214	GACTACTTACTCTGCACTTAGAACACATAGACAGAGGCGCTTGGAGGCGCATTTGACTTTGGG	1273
QY	1186	CCGGGCGCGCGCTGGGGATG 1204	
Db	1274	ACAGGTAAGGATGTGATG 1292	
RESULT 3			
US-09-764-864-21			
Sequence 21, Application US/09764864			
Patent No. US2002132753A1			
GENERAL INFORMATION:			
APPLICANT: Rosen et al.			
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies			
FILE REFERENCE: PT23			
CURRENT APPLICATION NUMBER: US/09/764,864			
CURRENT FILING DATE: 2001-01-17			
Prior application data removed - consult PALM or file wrapper			
NUMBER OF SEQ ID NOS: 1792			
SOFTWARE: PatentIn Ver. 2.0			
SEQ ID NO 21			
LENGTH: 1781			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-09-764-864-21			

Query Match	33.5%	Score 478.8	DB 10	Length 1781
Best Local Similarity	66.4%	Pred. No 6	7e-106	
Matches	703	Conservative	0	Mismatches 352; Indels 3; Gaps 17;
QY	226	CTGCTAGGGGATGCGCACATGACAACTTGGAGAGACAGCTATATTGGCCCATCTGC	285	
Db	137	CTGATCCAGAGAGGGGAAATCCCATGGAGAACTTGGAGAAACACACTGATCGCCCTATCTGC	196	
QY	286	CTGGAGATGTTTCCCAAGCCCGTGATCTTGGCCTGCGCAACACAACTGTGCGCGAAG	345	
Db	197	CTGGAGATGTTTACCACAGCCAGTGGTCATCTTGGCGGTGCCACACAACTGTGCGCGAAG	256	
QY	346	TGTGCCAAGACGCTCTTCCAGGCGCTCTATCCTCTGTGGCAATCCCGGGCTTCCAAAG	405	
Db	257	TGTGCCAAGACATCTTCCAGGCTGAAAAATCCCTACTGGACCAAGCCGGGGCAAGCTAGTG	316	
QY	406	GTCCTCTGAGAGAGCGTTCCGATGCCCATCTTGTAGGACAGAGGTTCTCCGTGCAGG	465	

Db	317	TCCATGCTCGGAGGCGCCGTTTCCGCTGCCGCCACCTGGCCGCCACGAGGATGATCATGATCGT	376
QY	466	CATGCTGCTATGAGCCCTGACACGGACACCTGCTATGTGGAGCAATCATTTGACATCTACAG	525
Db	377	CACGAGATTTCAGGCGCTGACAGGAACTGCTGGTGGGAACATCATTCGACATCTACAA	436
QY	526	CAGGAGTCTCCCGGCACTGCACGCCCAAGGCTGACAGCACTCATGTGTGAGGAGCAC	585
Db	437	CAGGAGTCTCCAGTCCGGCGCTGAGAAAGGAGT--CACCCCATGTGCAGAGAGCAC	493
QY	586	GAGGACGAGAAATACATCTACTGCTTGAGCTGACGAGGTGGCCCACTGCTCTCTGCG	645
Db	494	GAAAGTGAAGAAATTAACATCTACTGCTTCACGHTGAGGTGGCCACCTGCTCATGTGCG	553
QY	646	AAGGTTTCGGCGGCCCAAGAGACGTGAGGCGGCCCTGCGCCACCATTTACAAAGC	705
Db	554	AAGGTGTTGGGATCCACAAAGCCCTGGCAGGTGGCCCATTTGCAGAGTCTTCCAGGGA	613
QY	706	CAGAAAGTGAAGCTGAGCGATGGCATCGCATGCTGTGTGGCGGCAATGACCTGTGACG	765
Db	614	CAAAAGCTGAACATGAATAACTGTATCTTCATAGCTGTGTGGCGGGAATGACCGTGTGCG	673
QY	766	GCAGTGATCCACCGAATGGAGAGAGGTGTGCCAACCATTGGAGGAACAAGCGCGACAG	825
Db	674	ACCATCATCTCCTCAGCTGGAGGATTCGCCGTGAGTGACCAAGGAACAAGTACCAGTA	733
QY	826	AAGCACTGTTAAACGACAGGCTGCAGACCCCTGTGCGGGTTTGGAGAGCGCAAGGCG	885
Db	734	AAGGAAGAGCTGAGCCACAAGATTGTACACGTTGTATGTCATCTGTGATGAGAAGAAAGT	793
QY	886	GAACTGTTCAAGCACTGCGCCCGGGAGCAGAGAGAAAGTTGCAGCGCTCGGGGCGCTC	945
Db	794	GAGTGTCTGACGCGGATCACGAGGAGCAGAGAAAGAAAGCTTAGCTTATCGTAGGCGCTC	853
QY	946	ATTCGCCAGTACGAGACCATCTTGGAGGGCTCTCAAAAGCTGGTGGAGCGCCCAATCCAG	1005
Db	854	ATCCAGAGCTACAGAGAGAGCTGTGACAAAGTCCACAAAGCTGTGTGAAGTCCATCCAG	913
QY	1006	TCCATGGAGGAGCCCGACAGATGGCTTCTCACTCCAGCAGGCAAAAGAGCTGATCAACAAG	1065
Db	914	TCCCTGAGAGACCTGTGGGGGACCAACCTTCTTGTAGCTGCCCAACATCATCAAAAGC	973
QY	1066	GTCGGGGCATATCCAAAGTGTGAGCTGGCAGAGACGCGCGGAGCCAGGCTATGAGAGCATG	1125
Db	974	ATTGTGGAAGCTTCACAGGGCTGCCAGCTGGGGGAAGACAGACAGGCTTGGAGACATG	1033
QY	1126	GAGCAATTCCTGTGAGCGCTGAGACACGCGGGCCGAATTTCTCGAACCATCATCTCCAG	1185
Db	1034	GACTTCTTACTTTTGATTTTAAAGCACAATAGCAGAGCCCTTGATGAGCATTTACTTTGGG	1093
QY	1186	CCGGGCGCCGCTGGGGATGAAAGAGATGACGACATGGCTTTGGATGGGAGGAGGCAAT	1245
Db	1094	ACAGATGAGAGAGAGAAATTCATTGTGAAGAAAGATCAGAGAGAGAGAGAGATCCACA	1153
QY	1246	GCGGGGCTTGAAGAGAGGCGGCTGGACGTGCCAGAAAG	1283
Db	1154	GAAAGGAGAGAGAGACACCAAGTAAGAGAGCTGGATG	1191

RESULT 4  
 US-09-908-988B-3  
 ; Sequence 3, Application US/09908988B  
 ; Patent No. US20020127690A1  
 ;  
 GENERAL INFORMATION:  
 ;  
 ; APPLICANT: OLSON, ERIC  
 ;  
 ; APPLICANT: SPENCER, JEFFREY A.  
 ;  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES  
 ;  
 ; TITLE OF INVENTION: IN STRIATED MUSCLE CELLS  
 ;  
 ; FILE REFERENCE: MYOG:02805  
 ;  
 ; CURRENT APPLICATION NUMBER: US/09/908,988B  
 ;  
 ; CURRENT FILING DATE: 2000-07-18  
 ;  
 ; PRIOR APPLICATION NUMBER: 60/219,020  
 ;  
 ; PRIOR FILING DATE: 2000-07-18



```

QY 666 GGACTGTAGTGGCCCTCTGCCCCACATTTCACAAAGCCAGAAAGTGAAGTGAAGCA 725
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 619 AGACTGCCAGTGGTCCCTCCATCATATGTTCCAGAGACAGAAATGTGATCGATGA 678
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 726 TGGCATGCCAGTGGTGGGGGCAATGACCGTGTGACAGTATATACCCAGATGGA 785
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 679 TGGCATGCCATCTCTGTTGGGAGCAGACGATGATGCCAGGATGATACGACGCTGA 738
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 786 GGAGTGTGCCAGACCATTTAGAGACACGCCGACAGAAAGCACTGTAAACAGAG 845
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 739 AGACACCTGCAAAACTATACAGAGATGTTGAGAAAACAGAAACAGAGCTTTGTAGAA 798
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 846 GTTGAGACCTCTGTCGGGTTTGGAGAGCGCAAGGCGAACTCTTCAACACTGCC 905
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 799 GTTGATTAACCTGTATGATGAGAGAGAGAAATGAATGACCCCAAGTCATTTAC 858
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 906 CCGGAGAGAGAGAGAGATTCAGAGCGCTGCGGGGCTCATCCGCCATGACGAGACCA 965
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 859 CCGAACCAGAGAGAGAAATGGAACATGTCCGTCTGATCAAAAAGATTTCTGATCA 918
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 966 CTTCGAGGCTCTCAAACTGTGTGAGTCCGCCATCCATCCATGAGAGAGCCGAGAT 1025
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 919 TTTGAGAAACGTCTCAAAAGTTGATGAGCAATTCAGTTATGATGAGCCAGAAAT 978
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1026 GGCTCTTACCTCCAGCAGAGCAAGAGAGTGAATCAACAAAGTGGGGCAATGTCGAAGT 1085
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 979 GGCAAGTGTCTTGCAAAATGCCAAACCTGCTAAAAAAATCTCAGAAACATCAAAAGGC 1038
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1086 GGAGCTGGCAGAGAGCGCCGAGCCAGCATATGAGACATGAGCAATTCCTGTGAGCGT 1145
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1039 ATTTCAGATGAGAGAAATATGAAACATGTGATGAGAAATGAAACACTTACAGCAACCT 1098
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1146 GGAGCAGCTGCGCGAAATGTTGCAACATTCAGCTTCACCCGCGCGCTGGGATGA 1205
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1099 CAATAGAGAGAAAGATATACGTGAATTTGACTTTTACAGAGAGATGAAGATGAAGA 1158
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1206 AAGAGATGCGCATGCGCTTTGGATGGGAGAGCAATGCGGGGCTGGAGAGAGCG 1265
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1159 AAGAGAGAGAGCG--GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCA 1216
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1266 GCTGGA 1271
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1217 GTAGAA 1222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 6  
US-09-764-864-493  
; Sequence 493, Application US/09764864  
; Patent No. US20020132753A1

GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: Patent Ver. 2.0  
; SEQ ID NO 493  
; LENGTH: 587  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-864-493

Query Match 16.0%; Score 229.2; DB 10; Length 587;  
Best Local Similarity 70.9%; Pred. No. 3.8e-46;  
Matches 333; Conservative 1; Mismatches 129; Indels 7; Gaps 2;

```

QY 187 GAGCGCGGCGGATGAACTTCACGCGTTTCAAGCCGCTGTAGGGAGTGGCAGACAC 246
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124 GACAGCGAGAGATRACGCGCATCTGAAATTTACAAATCTTTTCCAAAGAGCAGCAGCC 183
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 247 ATGACAACTTGGAGAGACAGCTATTGCCCCATCTGCTGGAATGTCTCCAAAGCCC 306
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 ATGATTAACCTTAGAGAAACCACTCATCTGTCCCATCTGTAGAAATGTCTCAAGAACT 243
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 GTGTGATCTTGCCCTGCAACACACACACCTGTGCGGAGTGTGCCAGACGCTTTCCAG 366
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 244 GTGTGATCTTGCCCTGTCACACACACACCTGTGTAGGAATGTGCCAGATATTTTCCAG 303
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 367 GCCTTAATCTCTGTGGCAATCCCGGGGCTCCACAAGGTGTCTTCCAGAGAGCTTTC 426
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 304 GCTCTTAACCCGTATTGTCACCAAGAGAGAGTACACCATGCAATCAGGGGCGCATTC 363
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 427 CGATGCCATCTTGTAGCAGAGAGTGTCTGTCAGAGAGATGTGTATGCGCTGAC 486
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 364 CGTCCCATCTCTGTAGACATGAAATGTTTGTGATAGCATAGGGGTATATGACTTCAG 423
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 487 CGAAGCTGTAGTGGAGAACATCATGACATCTACAAAGAGAGTCTCCGGCCACTG 546
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 424 AGGAACCTGTGTGGAGAAATATCATTTGACATCTACAAAGCAGAGTCCACAGGCCA--- 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 547 CACGCCAGGCTGAACAGCACTCATGTGTGAGAGCAGAGAGAGAGAAATCAACATC 606
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 --GAAAGAAATCCGACACGCGCATGTGCGAGAGACATGAAGAGAGCGCATCAACATC 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 607 TACTGCTGAGCTGCGAGGTGCCACCTGCTCTCTGCAAGGTTTTCG 656
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 538 TACTGTCTGACGTGCGA--ATACCACTGCTCTGTGTGCAAGTGTTCG 586
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 7  
US-09-764-864-482  
; Sequence 482, Application US/09764864  
; Patent No. US20020132753A1

GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: Patent Ver. 2.0  
; SEQ ID NO 482  
; LENGTH: 446  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (362)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (400)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-864-482

Query Match 11.3%; Score 162; DB 10; Length 446;  
Best Local Similarity 72.5%; Pred. No. 4.7e-30;  
Matches 222; Conservative 0; Mismatches 82; Indels 2; Gaps 1;

```

QY 226 CTGCTAGGGGATGCGCACAACATGACACTTGGAGAGAGAGTCAATTTGCCCATCTGC 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 CTGATCCAGATGGGAATCCCATGAGAACTTGAGAAAGCATGATCTGCCATATCTGC 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 286 CTGAGATGTTCTCCAAAGCCGCTGTGATCTTGCCTCCCAACAACACTGTGCGCGAAG 345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 198 CTGGAATGTTTACCAAGCCAGTGTGATCTTGTGCGCGCAGCACAACCTGTGCGGAAG 257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 346 TGTGCCAAGAGCTCTTCCAGAGGCTTAAATCTGTGTGCAATCCGCGGCTCCACAAG 405
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 258 TGTGCCAATGACATCTTCCAGAGGCTGCAAAATCCTACTGACACACCGGCGCAGCTAGTG 317
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 406 GTGTCTTACAGAGAGAGTTCGATGCCCATCTTGTGAGACAGAGTGTGCTGACAGG 465
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 318 TCACATGCTGGAAGCCGTTTCGCTGCCCCACCTGCCGACGANGATGATGATGCT 377  
 QY 466 CATGCTGTATAGCCCTG--CAGCCGACCTGTAGTGAGACATCATGATCATGACA 523  
 Db 378 CACGAGGTGTAGGGCTGCGCAAMAGAACTGCTGGGTGGAACATCATCATCTACA 437  
 QY 524 AGCAGG 529  
 Db 438 AACAG 443

## RESULT 8

US-09-880-192-25  
 ; Sequence 25, Application US/09880192  
 ; Patent No. US2002007470A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Walker, Michael G.  
 ; APPLICANT: Volkmut, Wayne  
 ; APPLICANT: Klingler, Tod M.  
 ; APPLICANT: Azimzal, Yalda  
 ; TITLE OF INVENTION: POLYNUCLEOTIDES ASSOCIATED WITH CARDIAC MUSCLE FUNCTION  
 ; FILE REFERENCE: PB-0009-1 CIP  
 ; CURRENT APPLICATION NUMBER: US/09/880,192  
 ; NUMBER OF SEQ ID NOS: 62  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 25  
 ; LENGTH: 1039  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc-feature  
 ; OTHER INFORMATION: Incyte ID No. US2002007470A1 975169CT1  
 US-09-880-192-25

Query Match 9.9%; Score 141.6; DB 10; Length 1039;  
 Best Local Similarity 57.8%; Pred. No. 4.9e-25;

Matches 252; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

QY 848 TCAGACCTCTGCGCGGTTTGGAGAGCGCAAGGCGCAACTGCTTCAAGCACTGGCCC 907  
 Db 2 TTGACAGCTTATGCCATCTCTGATGAGAAAGATGATGCTGCAAGCGGATCAGC 61  
 QY 908 GGGAGCAGAGAGAGAGTTCAGCGCGCTGCGGCGCTCATCCGCAAGTACGGAGCACT 967  
 Db 62 AGGAGCAGAGAGAGAGTTCAGCTTATCAGAGCGCCCTCATCCAGAGTACAGAGCAAGC 121  
 QY 968 TGGAGGCTCTCAAGCTGCTGAGTCCGCGCTCATCCGCAAGTACGGAGCACT 967  
 Db 122 TGGACAGCTCACAAGCTGCTGAGTCCAGTCCATCCCTGAGCAGAGCTGGGGGAG 181  
 QY 1028 CTCCTTACCTCCAGCAGGCAAGAGAGTATCAACAAAGTCCGCAAGTACGGAGTGG 1087  
 Db 182 CCACCTTCTCTTATGACGCAAGCACTCATCAAAAGATGTTGGAAGCTTCAAGGGCT 241  
 QY 1088 AGCTGGCAGAGAGCGCCGAGCCAGCTATGAGAGCATGAGCAATTCCTGTGAGCGTGG 1147  
 Db 242 GCCAGCTGGGGAAACAGAGAGGCTTTGAGAACTGCTTCTTACTTGAATTAG 301  
 QY 1148 AGCAGCTGGCGGAATGTTGGCAACATGACTTCCAGCCGGGCGCCCTGGGGATGAAG 1207  
 Db 302 AGCAGATAGCAGAGCGCCCTGAGACCATTTGGAGCAGATAGAGAGAGAGAT 361  
 QY 1208 AGGATGAGCAGATGCTTTGGATGGGAGAGAGGCAATGCGGCTGGAGAGAGCGGC 1267  
 Db 362 TCATTGAAGAAGATGATCAGAGAGAGAGAGTCCACAGAGAGGAGAGAGAGACAC 421  
 QY 1268 TGACAGCGCCAGAGG 1283  
 Db 422 AGTAAGAGCTGATG 437

RESULT 9

US-09-969-852-4

; Sequence 4, Application US/09969852  
 ; Patent No. US2002013721A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Tianyan  
 ; APPLICANT: Liu, Huifen  
 ; APPLICANT: Li, Wei  
 ; APPLICANT: Zhao, Lipin  
 ; TITLE OF INVENTION: A METHOD FOR ESTABLISHING AN EXPRESSION SYSTEM OF SPIDER DRAGL  
 ; FILE REFERENCE: LIU-65  
 ; CURRENT APPLICATION NUMBER: US/09/969,852  
 ; PRIOR APPLICATION NUMBER: 2001-10-04  
 ; PRIOR FILING DATE: 2001-01-02  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 1852  
 ; TYPE: DNA  
 ; ORGANISM: Nephila clavipes  
 US-09-969-852-4

Query Match 4.2%; Score 59.6; DB 10; Length 1852;  
 Best Local Similarity 45.6%; Pred. No. 2.8e-05;  
 Matches 209; Conservative 0; Mismatches 249; Indels 0; Gaps 0;

QY 823 CAGAGCACTGTTAAACAGAGGTTGAGACCCCTGTCGCGGCTTTTGGAGAGCGCAAG 882  
 Db 257 CTGAGCAGCGCGCTGACAGCTGTGTGTCGCGGACAGAGAGATGAGAGTCTCGGA 316  
 QY 883 GGGCACTGCTTCAAGCAGTGGCCCGGAGCAGAGAGAGAGAGTTCACAGCGTGGGGCC 942  
 Db 317 GCCAAGTGCAGAGAGAGTGTAGTGGACAAAGGGGAGAGGAGTGCACAGCAGCAGCAG 376  
 QY 943 CTCATCCCGCAGTACGAGAGACCACTTGGAGGCTCTCAAGCTGTGTGAGTCCGCAATC 1002  
 Db 377 CTGAGAGTGCAGAGAGAGAGATGAGAGTCTTGGAGCCAGAGTCCAGAGCAGAGTGG 436  
 QY 1003 CACTCCATGAGAGAGCGCCAGATGCTCTTACTTCCAGCAGGCAAGAGAGTATCAAC 1062  
 Db 437 GATTAGTGGACAAAGGTGACAGGTGACAGCAGCAGCAGCAGCGAGGTCTGAGCAAG 496  
 QY 1063 AAGTCCGGGCAATGTGTGAGGTGAGCTGCGCAGAGAGCGCCGAGCCAGCTATGAGAGC 1122  
 Db 497 GAGGATACGTTGCTTGTGTGACAAAGTGCAGAGAGTGTGTATGAGAGCTTGGAA 556  
 QY 1123 ATGAGCAATTCCTGTGAGCTGTGAGCAGTGGCGCAAAATGTTGCGAAACATGACTTC 1182  
 Db 557 GCCAAGTGCAGAGCAGAGAGATTTAGTGGACAAAGTGTGACAGTGTGACAGCAGCAGCAG 616  
 QY 1183 CAGCCGGGCGCCCTGGGGATGAAGAGATGACAGCATGCTTTGGATGGGAGAGAGGC 1242  
 Db 617 CAGCTGAGGTGCGCGGCAAGAGAGAGTACTAGTGTGACAGAGTGTGCAAGAGAGAGTGGAG 676  
 QY 1243 AATCGGGGCTGGAGAGAGAGCGCTGAGCTGAGCTGACAGA 1280  
 Db 677 CAGCCCTGACAGCAGTGTGTGTCGCGACAGAGAGA 714

## RESULT 10

US-09-764-864-56  
 ; Sequence 56, Application US/09764864  
 ; Patent No. US20020132753A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PT223  
 ; CURRENT APPLICATION NUMBER: US/09/764,864  
 ; PRIOR APPLICATION data removed - consult PALM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 1792  
 ; SOFTWARE: PatentIn Ver. 2.0

```

; SEQ ID NO 56
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-56

```

Query Match	3.5%	Score 50.8;	DB 10;	Length 1400;
Best Local Similarity	51.3%;	Pred. No. 0.0033;		
Matches 118; Conservative	0;	Mismatches 112;	Indels 0;	Gaps 0;

QY 559 GAAAGACCTCAATGTTGTAGAGAGCAAGAGAGAGAAAGTATCAAACTCTACCTCGAG 618  
 Db 440 GACCCGGAGCCCAAGGTCTCGTGGTACACCCGAGAACCCGCTCACTTTCTCGAGAG 499  
 QY 619 TGCAGGTGCGCACTCTCTCTCTGCAAGTTTTCGGCCCCACAAGACTGTGAGTG 678  
 Db 500 GACCAAGAGCTCACTCTGTGGCTCTGGGGTCTGTTGGGCTTCCACCAACCAACCGGTC 559  
 QY 679 GCCCCTTGCCCACTATTTACAAAGCCCAAGAGTAGCTGAGCATTGGCATTCGGATG 738  
 Db 560 ACGCCGCTCTACACCGTCTGAGCCGATAAAGAGAGAGCTCCAGGCCCTTCTCTGAG 619  
 QY 739 CTGCTGGCGGGCAATGACCGTGTCCAGAGAGTGATTCACCAAGATGAGGA 788  
 Db 620 CTGAAGCAGAGCAAGAAAGAGTGTGATGAGCTCATCTGCCCAACTGGTGAA 669

RESULT 11  
US-09-880

```

US-09-880-192-41
Sequence 41, Application US/09880192
Patent No. US20020077470A1
GENERAL INFORMATION:
APPLICANT: Walker, Michael G.
APPLICANT: Volkmuth, Wayne
APPLICANT: Klingler, Tod M.
APPLICANT: Azimzai, Yalda
TITLE OF INVENTION: POLYNUCLEOTIDES ASSOCIATED WITH CARDIAC MUSCLE FUNCTION
FILE REFERENCE: PB-0009-1 CIP
CURRENT APPLICATION NUMBER: US/09/880,192
CURRENT FILING DATE: 2001-06-12
NUMBER OF SEQ ID NOS: 62
SOFTWARE: PERL Program
SEQ ID NO 41
LENGTH: 3441
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID NO. US20020077470A1 4151935CB1
US-09-880-192-41

```

Query Match	3.5%;	Score 50.4;	DB 10;	Length 3441;
Best Local Similarity	44.3%;	Pred. No. 0.0055;		
Matches 204;	Conservative	0;	Mismatches 256;	Indels 0;
			Gaps	0;

OY	578	AAACGACACGAGACGAGAGATCAACATCTACACGCTGAGCTGGGAGAGTGGCCACACGCT	637
Db	1088	AAGAGGTGAAAAAGTCCCAATTGACACATACATCTGTTACACTGCAAAATGTCCAATTTCTG	1147
OY	638	CTCTCTGCAAGGTTTTGGCGCCGCCACAAGCAGCTGTGAGTGGCCCCCTCTGCCACCAATTT	697
Db	1148	CCACTGACACAGGTTGTTGGACCCACCAAAACCATGAAATTTCAACCTTGGACACAGCTA	1207
OY	698	ACAAACGCGCAGAGAGTGTGAGCTGAGCGATGCGGATGCTGCTGGCGGGCAGATGACC	757
Db	1208	TAAAGTCTTAAAGTTTCATATTGACGAATTTTGTAGAAAAATTTTCAAGAAAAAGTCCTTGA	1267
OY	758	GTTGTGACGACGACGATCACCACCAATGTGAGAGAGGTTGCCACCATGAGAGACAACACCC	817
Db	1268	GGATTGAACCTTTGTTAGTGAATGTAAATCCTTTTAAATACCAATTGAGGAAAACTGTA	1327
OY	818	GCAGACAGAGCAACTGTTTAAACGACAGGTTTCGAGACCCCTGTGGCGGGTTTTTGAGAGAGC	877

Db	1328	GTAAAAATAGAAAAGGCTTAGAAGAACAGATATGAGAAATGATGAAACAAAGSTTTTAACAC	1387
Qy	878	GCAAGGCGCAACTGCTTCAAGCAGCTGCCCCGGAGCAGAGAGAGTGTGCAGCGCGCTGC	937
Db	1388	AGTATGATGAGAAAGCCCGAGCGCTTTGAGAGAGAGAAAGAAAGAAAGATGAGTGTCCGCG	1447
Qy	938	GGGGCCTCATCCGCCAGTACGAGACCACCTTGAGAGGCGCTCCTCAAGCTGTGGAGTCCG	997
Db	1448	ATGAGCAGATGGTCCACTTTTTCGACAGACATGGACACACTGCCMAAGACACCTTGGAGACCA	1507
Qy	998	CCATCCAGTCCATGAGAGAGCCGACAGATGGCTCTCTACT	1037
Db	1508	TGCTGAGAGAGACGAGAGACTTGTATAGGCCGCGCTTCCCT	1547

RESULT 12  
US-09-991

```

US-09-991-496-115
; Sequence 115, Application US/09991496
; Patent No. US20020169285A1
;
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Webb, John R.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Bhalla, Ajay
; APPLICANT: Coler, Rhea
; APPLICANT: Probst, Peter
; APPLICANT: Brannon, Mark
;
; TITLE OF INVENTION: LEISHMANIA ANTIGENS FOR USE IN THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF LEISHMANIASIS
; FILE REFERENCE: 210121.420C9
; CURRENT APPLICATION NUMBER: US/09/991,496
; CURRENT FILING DATE: 2001-11-20
; NUMBER OF SEQ. ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115
; LENGTH: 7065
; TYPE: DNA
; ORGANISM: Leishmania major and chagasti
US-09-991-496-115

```

Query Match	3.4%;	Score 48;	DB 9;	Length 7065;
Best Local Similarity	46.9%;	Pred. No. 0.026;		
Matches 254;	Conservative	0;	Mismatches 280;	Indels 8;
				Gaps 3,

QY	846	GTTCGAGACCCCTTGTGCGGCTTTTGGAGAGGCGCGAAGGGGAACTGCTTCAAGACACTGGC	905
Db	2190	GTCTGAGGGCCACAGGTGGCCAGCGCTGGCGCGGAGCCCGACAGGCGGGCCACGACGTGGC	2249
QY	906	CCGGGAGCAGAGGAGAGAAATTGCAAGCCCTGTGCGGGGCTCATCCGCCATGACGAGACCA	965
Db	2250	CGGGAACGCCGAGGAGGTGCACGACGCGCTTAGACACCGCCACGACGAGCAGCGCGCGAGCT	2309
QY	966	CTTGGAGGCTCTCTCAAAGCTGTGAGTCCGCCATCCAGTC--CATGAGAGACCGCA	10222
Db	2310	GGAAGCACACTTGGACAGCGCTGGCCGGCGGACCGGAGCGAGCGCGCCACGCTGTGGCGC	2369
QY	1023	GATGGCTCTACTCTCAGACGAGCAAGGAAAGGAGCTATCAACAAGGTGGGGCAATGTGAA	10822
Db	2370	GAACGCCGAGGAGCTGTGACGACCCCTAGACACCGCCACGACGAGCAGCGCGCCGAGCTGGA	24289
QY	1083	GGTGGACCTGGCAGAGACGCGCGGAGCCACGAGCTATGAGACGATGAGCAATTCATGTGAG	11442
Db	2430	GGCACAGGTGGCAAGCGCTGGCGGGGAGACCGGACGAGAGGCGGCCACGACGTGGCGCGGAA	24889
QY	1143	CGTGGACACATGGCGGAAATGTTGGCAACCATGACTTCCAGCCGGGCGCCGCTGGGGGA	12022
Db	2490	CGCGGAGGACTG---CAGCAGCGCTTAGACACCGCCACGACGAGCAGCGCGCGCGAGCTGGA	2546
QY	1203	TGGAAGAGATGACACAAATGGCTTTGGAAATGGGAGAGAGGGCAATCGGGGGCTGGAGGAGGA	12622
Db	2547	GGCCACAGGTGGCACGCGTGGCGCGGACCGCGACGAGAGGCGCCGCCACGACAGCTGGCGCGGAA	2606





```
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 54
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: -
; LOCATION: (1)-(1464)
; OTHER INFORMATION: "n" can be any nucleotide 'a', 'c', 'g' or 't'
US-09-771-161A-54

Query Match          3.1%; Score 43.8; DB 10; Length 1464;
Best Local Similarity 44.7%; Pred. No. 0.16;
Matches 263; Conservative 0; Mismatches 317; Indels 9; Gaps 2;
```

```
OY 442 AGGACGAGGTTGCTCTGGACGAGCATGTGTATGAGCCTGACGGGAACTGTAGTG 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 460 AGCCGCTGGCGGCTGAGAGAGCATGTCATGTAGCTGAGAGTGAACAGGGGGCG 519

OY 502 GAGAACATATGATGACATCTTACAGAGAGATCCTCCGGCCATCGACGCCAAGCTGAA 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 520 GAGAACATATGATGACATCTTACAGAGAGATCCTCCGGCCATCGAGTGA 579

OY 562 CAGCACTCATGTGTGAGAGAGAGAGAGAGAGAGATCAATCTACTGCTGAGCTGC 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 580 GCCCAGCAGATGTGGAGAGAGAGAGAGAGATGATGATCATCTCCGATCACTCCGC 639

OY 622 GAGGTGCCCACTGCTCTCTCTGCAAGGTTTGGCGCCCAAGAGACTGTGAGTGGCC 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 640 CGGGGCTGACAGCGCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 699

OY 682 CCTGTGCCCACTATTTACAAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGTG 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 700 CTTGACCTTGGGGGCTGTGAGAGCTGCGCATGAAAGAGCTGGCGCACCACTTCCAGTGGAG 759

OY 742 GTGGCGGCAATGACCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT---GTGCCAG 798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 760 CACGGGCTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 819

OY 799 ACCATTGAGAGACAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 852
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 820 GACCCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 879

OY 853 ACCCTGTGGCGGTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 912
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 880 CTGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 939

OY 913 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 972
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 940 GGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 999

OY 973 GGGCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1000 GGGCCCTTCCCGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1048
```

RESULT 18  
US-09-771-161A-53  
; Sequence 53, Application US/09771161A  
; Patent No. US20020110811A1

```
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53
; LENGTH: 2518
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: -
; LOCATION: (1)-(2518)
; OTHER INFORMATION: "n" can be any nucleotide 'a', 'c', 'g' or 't'
US-09-771-161A-53

Query Match          3.1%; Score 43.8; DB 10; Length 2518;
Best Local Similarity 44.7%; Pred. No. 0.15;
Matches 263; Conservative 0; Mismatches 317; Indels 9; Gaps 2;
```

```
OY 442 AGGACGAGGTTGCTCTGGACGAGCATGTGTATGAGCCTGACGGGAACTGTAGTG 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 460 AGCCGCTGGCGGCTGAGAGAGCATGTCATGTAGCTGAGAGTGAACAGGGGGCG 519

OY 502 GAGAACATATGATGACATCTTACAGAGAGATCCTCCGGCCATCGACGCCAAGCTGAA 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 520 GAGAACATATGATGACATCTTACAGAGAGATCCTCCGGCCATCGAGTGA 579

OY 562 CAGCACTCATGTGTGAGAGAGAGAGAGAGAGAGATCAATCTACTGCTGAGCTGC 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 580 GCCCAGCAGATGTGGAGAGAGAGAGAGAGAGATGATGATCATCTCCGATCACTCCGC 639

OY 622 GAGGTGCCCACTGCTCTCTCTGCAAGGTTTGGCGCCCAAGAGACTGTGAGTGGCC 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 640 CGGGGCTGACAGCGCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 699

OY 682 CCTGTGCCCACTATTTACAAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGTG 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 700 CTTGACCTTGGGGGCTGTGAGAGCTGCGCATGAAAGAGCTGGCGCACCACTTCCAGTGGAG 759

OY 742 GTGGCGGCAATGACCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTGTG 798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 760 CACGGGCTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 819

OY 799 ACCATTGAGAGACAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 852
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 820 GACCCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 879

OY 853 ACCCTGTGGCGGTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 912
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 880 CTGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 939

OY 913 CAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 972
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 940 GGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 999

OY 973 GGGCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1000 GGGCCCTTCCCGCAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1048
```

RESULT 19  
US-09-771-208-20/c  
; Sequence 20, Application US/09771208  
; Patent No. US2002015564A1  
; GENERAL INFORMATION:  
; APPLICANT: MEDRANO, JUAN  
; APPLICANT: BRADFORD, ERIC  
; APPLICANT: HORVAT, SIMON







GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 10:09:37 ; Search time 2857 Seconds

(Without alignments)  
14576.865 Million cell updates/sec

Title: US-09-908-988b-1

Sequence: 1 aagaggtgtagacagagtggt.....ataaagactcaagtgctccc 1431

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 13

Total number of hits satisfying chosen parameters: 36437

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

GenBank:\*

1: gb\_ba:\*

2: gb\_ba:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vi:\*

30: em\_hum:\*

31: em\_hum:\*

32: em\_hum:\*

33: em\_hum:\*

34: em\_hum:\*

35: em\_hum:\*

36: em\_hum:\*

37: em\_hum:\*

38: em\_hum:\*

39: em\_hum:\*

40: em\_hum:\*

41: em\_hum:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1431	100.0	1431	6 AX418848	AX418848 Sequence
2	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
3	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
4	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
5	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
6	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
7	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
8	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
9	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
10	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
11	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
12	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
13	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
14	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
15	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
16	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
17	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
18	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
19	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
20	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
21	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
22	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
23	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
24	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
25	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
26	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
27	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
28	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
29	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
30	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
31	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
32	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
33	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
34	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
35	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
36	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
37	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
38	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
39	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
40	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
41	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
42	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
43	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
44	1431	100.0	1448	10 AF294790	AF294790 Mus muscu
45	1431	100.0	1448	10 AF294790	AF294790 Mus muscu

## ALIGNMENTS

RESULT 1	AX418848	LOCUS	DEFINITION	AX418848	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
Sequence 1	from Patent WO0206318.								
AX418848.1	GI:21523712								
house mouse.									
Mus musculus									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.									

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0206318-A 1 24-JAN-2002;  
BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)  
FEATURES Location/Qualifiers  
source 1.1431  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
199.1299  
/note="unnamed protein product"  
/protein\_id="CAD35455.1"  
/db\_xref="GI:21523713"  
/translation="MNFTVGFKPLIDGHNMDNEKOLICPIOLEMFKPVVILPCQH  
MLCRKANDVQASNPIMQSRSTTVSSGSRFRCSREHVVLDHGVGLQRLNLYE  
NIIDYKQESSRPLHAKAEQIHMCBEHDEKINIVCLSEVPTCSIKYFGAHKDEE  
APLPTIYKRSKSELSDGILAMLVAGNDRVQAVITOMEVQOUTIEDNSRROKOLLNORE  
TICVALEERGELLQALAREQEKIQRVGLIROYGDHLEGGSKIVESAIOSEEPOM  
ALVIOAKELINKYKAMSKYELAGRPERYSEMEPFSVSEVHAEMLRITIDPOGAAAG  
DEEDDMALDGEBSNAGLEERLDVPEBSGLH"

CDS

BASE COUNT 338 a 384 c 473 g 236 t  
ORIGIN

Query Match 100.0%; Score 1431; DB 6; Length 1431;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AAGAGGTGTAGACAGAGTGTCTGGAATATAGACAGGGGTGAGAGAGCTGTAGGGAGAG 60  
1 AAGAGGTGTAGACAGAGTGTCTGGAATATAGACAGGGGTGAGAGAGCTGTAGGGAGAG 60  
61 GACAGGACTCTTCCAAAGGAGGAGCAATATGCCGGGATCCCAAGATCAGTCAAGCTTAAC 120  
61 GACAGGACTCTTCCAAAGGAGGAGCAATATGCCGGGATCCCAAGATCAGTCAAGCTTAAC 120  
121 TGACCGAGAGAGGTGACACAGGAGGAGGAGCAAGCAAGAGGAGGAGGAGGAGGAGGAG 180  
121 TGACCGAGAGAGGTGACACAGGAGGAGGAGCAAGCAAGAGGAGGAGGAGGAGGAGGAG 180  
121 TGACCGAGAGAGGTGACACAGGAGGAGGAGCAAGCAAGAGGAGGAGGAGGAGGAGGAG 180  
181 GCTCCAGAGCGCGCGGGGAGTGAATCTCAAGGTGTTTCAAGCCGCTGTAGGGATGCG 240  
181 GCTCCAGAGCGCGCGGGGAGTGAATCTCAAGGTGTTTCAAGCCGCTGTAGGGATGCG 240  
181 GCTCCAGAGCGCGCGGGGAGTGAATCTCAAGGTGTTTCAAGCCGCTGTAGGGATGCG 240  
181 GCTCCAGAGCGCGCGGGGAGTGAATCTCAAGGTGTTTCAAGCCGCTGTAGGGATGCG 240  
241 CACACATGAGCAACTGGAGAGAGAGCTATTTGCCCCATGTGCTGGAGATGTTCTCC 300  
241 CACACATGAGCAACTGGAGAGAGAGCTATTTGCCCCATGTGCTGGAGATGTTCTCC 300  
241 CACACATGAGCAACTGGAGAGAGAGCTATTTGCCCCATGTGCTGGAGATGTTCTCC 300  
301 AAGCCCGTGTGATCTTGCCTTGCCCAACACAACTGTGCGCAAGTGTGCCAAGCACTC 360  
301 AAGCCCGTGTGATCTTGCCTTGCCCAACACAACTGTGCGCAAGTGTGCCAAGCACTC 360  
361 TTCAGAGCTCTTAATCTCTGTGTGCAATCCCGGGGCTCCACAGGGTGTCTTCAGAGAGA 420  
361 TTCAGAGCTCTTAATCTCTGTGTGCAATCCCGGGGCTCCACAGGGTGTCTTCAGAGAGA 420  
421 CGTTCCGATGCCATCTTGTGAGCAGAGGTTGTCTGAGACAGGAGCATGTCTATATGTC 480  
421 CGTTCCGATGCCATCTTGTGAGCAGAGGTTGTCTGAGACAGGAGCATGTCTATATGTC 480  
481 CTGACGCGGAACCTGTAGTGAAGACATCTTGCATCTAACAGAGAGAGTCTCCCG 540  
481 CTGACGCGGAACCTGTAGTGAAGACATCTTGCATCTAACAGAGAGAGTCTCCCG 540  
481 CTGACGCGGAACCTGTAGTGAAGACATCTTGCATCTAACAGAGAGAGTCTCCCG 540  
481 CTGACGCGGAACCTGTAGTGAAGACATCTTGCATCTAACAGAGAGAGTCTCCCG 540  
541 CCAGCTGACGCGGAACCTGTAGTGAAGACATCTTGCATCTAACAGAGAGAGTCTCCCG 600  
541 CCAGCTGACGCGGAACCTGTAGTGAAGACATCTTGCATCTAACAGAGAGAGTCTCCCG 600  
601 AACATCTAGTCCCTGAGCTGAGAGTGGCCCACTGTCTCTCTGCAAGGTTTTCGGCGCC 660  
601 AACATCTAGTCCCTGAGCTGAGAGTGGCCCACTGTCTCTCTCTGCAAGGTTTTCGGCGCC 660  
661 CACAGAGAGTGTAGTGGCGCCCTCTGCGCCACCATTTTCAAAAGCCAGAGAGTGAAGT 720  
661 CACAGAGAGTGTAGTGGCGCCCTCTGCGCCACCATTTTCAAAAGCCAGAGAGTGAAGT 720

QY 721 AGCGATGCGATCGCGATGCTGTGGCGGCAATGACCGTGTGCGAGGATGATCACCAG 780  
DB 721 AGCGATGCGATCGCGATGCTGTGGCGGCAATGACCGTGTGCGAGGATGATCACCAG 780  
QY 781 ATGAGAGAGGTGTGCCAGACCATTTAGAGACACAGCCGACAGAGAGCAACTGTTAAAC 840  
DB 781 ATGAGAGAGGTGTGCCAGACCATTTAGAGACACAGCCGACAGAGAGCAACTGTTAAAC 840  
QY 841 CAGAGGTTGCAACCTGTGGCGGCTTTTGGAGAGAGGCGCAAGGCGCAACTGTTCAAGCA 900  
DB 841 CAGAGGTTGCAACCTGTGGCGGCTTTTGGAGAGAGGCGCAAGGCGCAACTGTTCAAGCA 900  
QY 901 CTGGCCCGGAGACAGAGAGAGAGAGTGTGACAGCGCGTGGCGGCGCATCCCGCATACGGA 960  
DB 901 CTGGCCCGGAGACAGAGAGAGAGAGTGTGACAGCGCGTGGCGGCGCATCCCGCATACGGA 960  
QY 961 GACCACTTGGAGAGGCTCTCTCAAAAGCTGTGTGAGTCCGCCATTCAGTCCATGAGAGAGCG 1020  
DB 961 GACCACTTGGAGAGGCTCTCTCAAAAGCTGTGTGAGTCCGCCATTCAGTCCATGAGAGAGCG 1020  
QY 1021 CAGATGGCTCTCTACCTCCAGAGGCAAGAGAGCTGATCAACAAAGTGGGGCAATGTCG 1080  
DB 1021 CAGATGGCTCTCTACCTCCAGAGGCAAGAGAGCTGATCAACAAAGTGGGGCAATGTCG 1080  
QY 1081 AAGTGAAGCTGTGACAGAGAGCGCGGAGCGAGCGCTATGAGAGAGTGAATCTCTGTG 1140  
DB 1081 AAGTGAAGCTGTGACAGAGAGCGCGGAGCGAGCGCTATGAGAGAGTGAATCTCTGTG 1140  
QY 1141 AGCGTGAAGACAGTGGCGCAATTTGCGAACCATGACACTTCCAGCCGGCGCGCTGTGG 1200  
DB 1141 AGCGTGAAGACAGTGGCGCAATTTGCGAACCATGACACTTCCAGCCGGCGCGCTGTGG 1200  
QY 1201 GATGAAGAGATGACAGACATGTGCTGTGATGGAGAGAGGAGGAGGAGGAGGAGGAG 1260  
DB 1201 GATGAAGAGATGACAGACATGTGCTGTGATGGAGAGAGGAGGAGGAGGAGGAGGAG 1260  
QY 1261 GAGCGCTGTGAGAGTGTGACAGAGGCTGAGGCTGACAGCCGACTGTGATCCAGAGCGC 1320  
DB 1261 GAGCGCTGTGAGAGTGTGACAGAGGCTGAGGCTGACAGCCGACTGTGATCCAGAGCGC 1320  
QY 1321 ACACCGGAAGCGGAGGAGGAGGATGCTGAGATCTGCGGAGAGACACCGCGCACCA 1380  
DB 1321 ACACCGGAAGCGGAGGAGGAGGATGCTGAGATCTGCGGAGAGACACCGCGCACCA 1380  
QY 1381 GCTGGCTTCCCGCGCGGAGAGGTTTCAATTAAGAGACTCAAGTGTCC 1431  
DB 1381 GCTGGCTTCCCGCGCGGAGAGGTTTCAATTAAGAGACTCAAGTGTCC 1431

RESULT 2  
AF294790 1448 bp mRNA linear ROD 08-SEP-2000  
LOCUS  
DEFINITION Mus musculus RING-finger protein MURF mRNA, complete cds.  
ACCESSION AF294790.1 GI:9945009  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1448)  
REFERENCE  
AUTHORS Spencer, J.A., Eliazzer, S., Ilarria, R.L. Jr., Richardson, J.A. and Olson, E.N.  
TITLE Regulation of microtubule dynamics and myogenic differentiation by MURF, a striated muscle RING-finger protein  
JOURNAL J. Cell Biol. 150 (4), 771-784 (2000)  
MEDLINE 10953002  
PUBMED 2 (bases 1 to 1448)  
REFERENCE  
AUTHORS Spencer, J.A. and Olson, E.N.  
TITLE Direct Submission  
JOURNAL Submitted (09-AUG-2000) Molecular Biology, UTSW, 6000 Harry Hines Blvd., Dallas, TX 75390-9148, USA

FEATURES	Location/Qualifiers
source	1. .1448

BASE COUNT	355 a	384 c	473 g	236 t
ORIGIN				

Query Match	100.0%;	Score 1431;	DB 10;	Length 1448;
Best Local Similarity	100.0%;	Pred. NO. 0;		
Matches 1431;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

QY	1	AAGCAGTGTAGACAGAGTGTCTGGAAATATGACAGGGGTGAGAGGAGCTGTTAGGGGAAG	60
Db	1	AAGAGTGTACACAGAGTGTCTGGAAATATGACAGGGGTGAGAGGAGCTGTTAGGGGAAG	60
QY	61	GACAGGACTCTTCCAAAGAGGAGCAATATGCCGGATCCCAAGATCCAGTCAAGCTTAAC	120
Db	61	GACAGGACTCTTCCAAAGAGGAGCAATATGCCGGATCCCAAGATCCAGTCAAGCTTAAC	120
QY	121	TGACCGAGAGGAGGGTGCACAGGCGGAGAAAGGCACAGACAGGGGCCACAGCGCGGAG	180
Db	121	TGACCGAGAGGAGGGTGCACAGGCGGAGAAAGGCACACAGAGGGGCCACACAGCAGGAG	180
QY	181	GCTCCAGAGCCGCGGGGATGAATCTCAGGTGGGTTTCAAGCCGCTGCTAGGGATCG	240
Db	181	GCTCCAGAGCCGCGGGGATGAATCTCAGGTGGGTTTCAAGCCGCTGCTAGGGATCG	240
QY	241	CACAAACATGGAACATTTGGAGAGAGCTCATTTGGCCCATCTGCTGTGAGATGTTTCC	300
Db	241	CACAAACATGGAACATTTGGAGAGAGCTCATTTGGCCCATCTGCTGTGAGATGTTTCC	300
QY	301	AAGCCCGTGTGATCTTCCCTGCCAACACACAACTGTGCCCAATGTGCCAACACAGTC	360
Db	301	AAGCCCGTGTGATCTTCCCTGCCAACACAACTGTGCCCAATGTGCCAACACAGTC	360
QY	361	TTCCAGGCTCTAATCTCTGTGTGGCAATCCGGGGCTCACAAAGGTCTTCAGGAGA	420
Db	361	TTCCAGGCTCTAATCTCTGTGTGGCAATCCGGGGCTCACAAAGGTCTTCAGGAGA	420
QY	421	CGTTCCGATGCCCATCTTGTAGAGCAGAGTGTCTCGAGACGGCATGTGTCTATGGC	480
Db	421	CGTTCCGATGCCCATCTTGTGTAGGACAGAGTGTCTCGAGACGGCATGTGTCTATGGC	480
QY	481	CTGCAGCGAACTCTCTAGTGGAGAACTAATTGACATCTACAAACAGAGTCTCCGG	540
Db	481	CTGCAGCGAACTCTCTAGTGGAGAACTAATTGACATCTACAAACAGAGTCTCCGG	540
QY	541	CCACTGCAGCGCCAAAGGCTGAACAGCACTCATGTGTGAGAGACAGAGGACGAGAATC	600
Db	541	CCACTGCAGCGCCAAAGGCTGAACAGCACTCATGTGTGAGAGACAGAGGACGAGAATC	600
QY	601	AACATCTACTGCTGAGTGGAGAGTGGCCACCTCTCTCTGTGAAGTTTTCGGCGC	660
Db	601	AACATCTACTGCTGAGTGGAGAGTGGCCACCTCTCTCTGTGAAGTTTTCGGCGC	660
QY	661	CACAAGGACTGTGAGTGGCCCTCTGGCCACATTTACAAAGCCAGAGAGTAGCTG	720
Db	661	CACAAGGACTGTGAGTGGCCCTCTGGCCACATTTACAAAGCCAGAGAGTAGCTG	720
QY	721	ACCGATGGCANTCGCATGCTGTGGCGGGGAATGACCGTGTCCAGGACGTATCAACCG	780

Db	721	AGCGATGGCATTCGGGATGCTGTGTGGCGGGCATGACCGTGTGCGGCATGATACCCAG	780
QY	781	ATGAGAGAGAGTGTGCCACAGACCATTGAGAGCAACAACGCCCAACAGAAAGCAACTGTTAAAC	840
Db	781	ATGGAGGAGGTGTGTCCAGAACCATTTGAGAGCAACAAGCCGACACAGAAAGCAACTGTTAAAC	840
QY	841	CAGAGGTTCCGAGACCTCTGTGCGCGGTTTTTGGAGAGCGCCAAAGGGCAACTCTCTTAAGCA	900
Db	841	CAGAGGTTCCGAGACCTGTGTGCGCGGTTTTTGGAGAGCGCCAAAGGGCGAACTCTTCAACCA	900
QY	901	CTGGCGCCGGAGACAGAGAGGAGATTGGTAGGCGCGGCGGGCCCTCATCGCCAGTAGCGA	960
Db	901	CTGGCGCCGGAGACAGAGAGGAGATTGGTAGGCGCGGTGCGGGGCTTCATCCGCACAGTAGCGA	960
QY	961	GACCACTTGGAGGGCTCCTCAAGAGCTGGTGGAGTCCGCGCATCCAGTCCATGGAGAGCCG	1020
Db	961	GACCACTTGGAGGGGCTCCTCAAGAGTGGTAGTCCGCGCATCCAGTCCATGGAGAGCCG	1020
QY	1021	CAGATGCGTCTCTTACTCTCCAGCAGGCAAAAGAGCTGATCAACAAAGTTCGGGGCAATGTGC	1080
Db	1021	CAGATGCGTCTCTTACTCTCCAGCAGGCAAAAGAGGCTGATCAACAAAGTTCGGGGCAATGTGC	1080
QY	1081	AAGGAGGAGCTGGGAGGAGCGCGCGGAGGCCAGTATGAGACATGGAGCAATTCCTGTG	1140
Db	1081	AAGGTGAGGCTGGCAGGAGCGCGCGGAGGCCAAGGCTTGTGAGACATGGAGCAATTCCTGTG	1140
QY	1141	AGCGTGGAGCAAGTGGCCGAATGTTGGCAACCATGACCTTCACAGCCGGGCGCGCTGGG	1200
Db	1141	AGCGTGGAGCAAGTGGCCGAATGTTGGCAACCATGACCTTCACAGCCGGGCGCGCTGGG	1200
QY	1201	GATGAAGAGGATGACGACATGCGTTTGGATGGGAGGAGAGGGCAATGCGGGGCTGGAGAG	1260
Db	1201	GATGAAGAGGATGAGCAATGCGTTTGGATGGGAGGAGAGGGCAATGCGGGGCTGGAGAGAG	1260
QY	1261	GAGCGGCTGGAGCTGCCAGAAAGGCTCAGAGCCTGCACCTAACCCGACTGTGATCCAGAGGCG	1320
Db	1261	GAGCGGCTGGAGCTGCCAGAAAGGCTCAGAGCCTGCACCTAACCCGACTGTGATCCAGAGGCG	1320
QY	1321	ACACCCGAAGCGGGAGCCAAAGGATGCTGAGAGATCTGCGCAGAGACCAACCGCGCCACCAA	1380
Db	1321	ACACCCGAAGCGGGAGCCAAAGGATGCTGAGAGATCTGCGCAGAGACCAACCGCGCCACCAA	1380
QY	1381	GCTGGGCTTCCCGCCCGCGGGAAGGTTCTCAATAAAGAGCTCAATGTGCC	1441
Db	1381	GCTGGGCTTCCCGCCCGCGGGAAGGTTCTCAATAAAGAGCTCAATGTGCC	1441

RESULT 3  
 AC084883  
 LOCUS  
 DEFINITION  
 AC084883  
 AC084883  
 AC084883  
 AC084883.10 GI:11976433  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
 Mus musculus.  
 Mus musculus.  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 209211)  
 Han,J., Montomey,K.T., Grills,G., Chiu,D., Decker,J., Fusina,M.,  
 Goltz,J., Halder,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J.,  
 Pereira,A., Shim,C., Thomas,E. and Kucherlapati,R.  
 High Throughput Mouse Sequencing  
 Unpublished  
 2 (bases 1 to 209211)  
 Han,J., Montomey,K.T., Grills,G., Chiu,D., Decker,J., Fusina,M.,  
 Goltz,J., Halder,A., Hall,L., Ioshikhes,I.P., Lee,E., Long,J.,  
 Pereira,A., Shim,C., Thomas,E. and Kucherlapati,R.  
 Direct Submission  
 Submitted (29-NOV-2000) Department of Molecular Genetics, Albert  
 Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,

## COMMENT

Bronx, NY 10461, USA  
On Dec 21, 2001 this sequence version replaced g1:15148085.

-----Genome Center

Center: Harvard Partners Genome Center

Center Code: HPGC

Web site: <http://www.hpgc.org/Sequence/mouse.html>  
Contact: [hpgc@emdel.mgh.harvard.edu](mailto:hpgc@emdel.mgh.harvard.edu)

-----Summary Statistics

Center project name: ABD

Sequencing vector: pUC18; 108752

Chemistry: Dye-terminator Big Dye; 100%

\*Consensus quality: 204023 at least Q20

\*Consensus quality: 203326 at least Q30

\*Consensus quality: 201766 at least Q40

\*Estimated insert size: agarose-PP - N/A

\*Estimated insert size: 208811 - sum-of-contigs

Quality coverage: agarose-PP - N/A

Quality coverage: 10.2 x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 21 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```

1      30156      30175: contig of 30155 bp in length
*      30176      30175: gap of unknown length
*      30176      76428: contig of 46253 bp in length
*      76429      76448: gap of unknown length
*      76449      99269: contig of 22821 bp in length
*      99270      99289: gap of unknown length
*      99290      119049: contig of 19740 bp in length
*      119030      119049: gap of unknown length
*      119050      133795: contig of 14746 bp in length
*      133796      133815: gap of unknown length
*      133816      147516: contig of 13701 bp in length
*      147517      147536: gap of unknown length
*      147537      161508: contig of 13972 bp in length
*      161509      161528: gap of unknown length
*      161529      171555: contig of 10027 bp in length
*      171556      171575: gap of unknown length
*      171576      182235: contig of 10660 bp in length
*      182236      182255: gap of unknown length
*      182256      190163: contig of 7908 bp in length
*      190164      190183: gap of unknown length
*      190184      194456: contig of 4273 bp in length
*      194457      194476: gap of unknown length
*      194477      199061: contig of 4585 bp in length
*      199062      199081: gap of unknown length
*      199082      201315: contig of 2234 bp in length
*      201316      201335: gap of unknown length
*      201336      201781: contig of 446 bp in length
*      201782      201801: gap of unknown length
*      201802      203287: contig of 1486 bp in length
*      203288      203307: gap of unknown length
*      203308      204141: contig of 834 bp in length
*      204142      204161: gap of unknown length
*      204162      204246: contig of 85 bp in length
*      204247      204266: gap of unknown length
*      204267      205695: contig of 1429 bp in length
*      205696      205715: gap of unknown length
*      205716      206923: contig of 1208 bp in length
*      206924      206943: gap of unknown length
*      206944      207861: contig of 898 bp in length
*      207862      207881: gap of unknown length
*      207882      209211: contig of 1350 bp in length.

```

## FEATURES

## SOURCE

1. .209211  
/strain="Mus musculus"  
/db\_xref="taxon:10090"

```

/misc_feature /chromosome="5"
/misc_feature /clone="RP23-154K6"
/misc_feature /sex="male"
1. .30155
/misc_feature /note="assembly_name:Contig39"
30176. .76428
/misc_feature /note="assembly_name:Contig38"
76449. .99269
/misc_feature /note="assembly_name:Contig37"
99290. .119029
/misc_feature /note="assembly_name:Contig36"
119050. .133795
/misc_feature /note="assembly_name:Contig35"
clone_end:T7
vector_side:right"
133816. .147516
/misc_feature /note="assembly_name:Contig34"
147537. .161508
/misc_feature /note="assembly_name:Contig33"
161529. .171555
/misc_feature /note="assembly_name:Contig32"
clone_end:SP6
vector_side:right"
171576. .182235
/misc_feature /note="assembly_name:Contig31"
182256. .190163
/misc_feature /note="assembly_name:Contig30"
190184. .194456
/misc_feature /note="assembly_name:Contig29"
194477. .199061
/misc_feature /note="assembly_name:Contig28"
199082. .201315
/misc_feature /note="assembly_name:Contig27"
201336. .201781
/misc_feature /note="assembly_name:Contig26"
201802. .203287
/misc_feature /note="assembly_name:Contig25"
203308. .204141
/misc_feature /note="assembly_name:Contig24"
204162. .204246
/misc_feature /note="assembly_name:Contig23"
204267. .205695
/misc_feature /note="assembly_name:Contig22"
205716. .206923
/misc_feature /note="assembly_name:Contig21"
206944. .207861
/misc_feature /note="assembly_name:Contig20"
207862. .209211
/misc_feature /note="assembly_name:Contig19"

BASE COUNT 55539 a 49732 c 50169 g 53288 t 483 others
ORIGIN

Query Match 25.6% Score 367; DB 2; Length 209211;
Best Local Similarity 100.0%; Pred. No. 5; 6e-192;
Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGGAGTGTAGACAGAGTGTCTGGAATATAGACAGGGTGAAGAGAGGCTTAAAGGGAAG 60
|||||
Db 190769 AAGGAGTGTAGACAGAGTGTCTGGAATATAGACAGGGTGAAGAGAGGCTTAAAGGGAAG 190828

Qy 61 GACAGACTCTTCCAGAGGAGCAATAGCCGGGATCCCAAGATCCAGTACGCTTAAC 120
|||||
Db 190829 GACAGACTCTTCCAGAGGAGCAATAGCCGGGATCCCAAGATCCAGTACGCTTAAC 190888

Qy 121 TGACCGAGGAAGGCTCCACAGCGAGGGGAAGAAGCCACAGCAGGCGCACAGGAG 180
|||||
Db 190889 TGACCGAGGAAGGCTCCACAGCGAGGGGAAGAAGCCACAGCAGGCGCACAGGAG 190948

Qy 181 GCTCCAGAGCGCGCGGAGTGAAGTTCACGCTGGGTTTCAAGCGGCTTGGGATCG 240
|||||
Db 190949 GCTCCAGAGCGCGCGGAGTGAAGTTCACGCTGGGTTTCAAGCGGCTTGGGATCG 191008

Qy 241 CACACATGAGACAATTGGAGAGAGCAAGCTTATTTGCCCATCTGCTGGAGATGTTTCC 300

```

QY	361	TTCCAGC	367	TTTTTTT	Db 191129	TTCCAGC	191135
Db 191009	CACACATGAGCAACACTTGGAGAGACGCTCATCTTGGCCCATCTGGCCATGCGATGTCGCC	191068					
QY	301	AAGCCCGTGGTATCTTGGCCCTGGCCACACACACTGTGGCCGAAGTGTGCCACGACGTC	360	TT			
Db 191069	AAGCCCGTGGTATCTTGGCCCTGGCCACACACACTGTGGCCGAAGTGTGCCACGACGTC	191128					
QY	361	TTCCAGC	367	TTTTTTT			
Db 191129	TTCCAGC	191135					
RESULT 4	AC114619	199296 bp	DNA	linear	HTG 06-JUN-2007		
LOCUS	AC114619						
DEFINITION	Mus musculus clone RP24-86123, WORKING DRAFT SEQUENCE, 15 ordered						
ACCESSION	AC114619						
VERSION	AC114619.2	GI:21327612					
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.						
SOURCE	house mouse.						
ORGANISM	Mus musculus						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus						
AUTHORS	1 (bases 1 to 199296) Bliren, B., Linton, L., Nusbaum, C. and Lander, E.						
TITLE	Mus musculus, clone RP24-86123						
JOURNAL	Unpublished						
REFERENCE	2 (bases 1 to 199296) Bliren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barra, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campoliano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, R., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Plerre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Mathews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riebeck, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliou, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.						
TITLE	Direct Submission						
JOURNAL	Submitted (10-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA						
REFERENCE	3 (bases 1 to 199296)						
AUTHORS	Bliren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barra, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campoliano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, R., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Plerre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Mathews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riebeck, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliou, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.						

```

TITLE
JOURNAL
COMMENT
Stojanovic,N., Straus,N., Subramanian,A., Palamas,J., Testa,E.S.,
Theodorou,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zairoun,J., Zembek,L., Zimmer,A. and Zody.M.
Direct Submission
Submitted (06-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 6, 2002 this sequence version replaced g1:193111148.
All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L24325
Center clone name: 86.L1.23
-----
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 193147 bases at least Q40
Consensus quality: 196185 bases at least Q30
Consensus quality: 197247 bases at least Q20
Insert size: 194000; agarose-fp
Insert size: 197896; sum-of-contigs
Quality coverage: 6.1 in Q20 bases; agarose-fp
Quality coverage: 6.0 in Q20 bases; sum-of-ctngs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1
* 728: contig of 728 bp in length
* 729 828: gap of 100 bp
* 829 2029: contig of 1201 bp in length
* 2030 2129: gap of 100 bp
* 2130 3803: contig of 1674 bp in length
* 3804 3903: gap of 100 bp
* 3904 6051: contig of 2148 bp in length
* 6052 6151: gap of 100 bp
* 6152 10082: contig of 3931 bp in length
* 10083 10182: gap of 100 bp
* 10183 18316: contig of 8134 bp in length
* 18317 18416: gap of 100 bp
* 18417 27114: contig of 8698 bp in length
* 27115 27214: gap of 100 bp
* 27215 33934: contig of 6720 bp in length
* 33935 34034: gap of 100 bp
* 34035 44466: contig of 10432 bp in length
* 44467 44566: gap of 100 bp
* 44567 59489: contig of 14923 bp in length
* 59490 59589: gap of 100 bp
* 59590 79430: contig of 19841 bp in length
* 79431 79530: gap of 100 bp
* 79531 99545: contig of 20015 bp in length
* 99546 99645: gap of 100 bp
* 99646 123390: contig of 23645 bp in length
* 123391 123390: gap of 100 bp
* 123391 151687: contig of 28297 bp in length
* 151688 151787: gap of 100 bp
* 151788 199296: contig of 47509 bp in length.
Location/Qualifiers
1. 199296
/organism="Mus musculus"
/db_xref="taxon:10090"

```

```

misc_feature /clone="RP24-86123"
              /clone.lib="RP24 Male Mouse BAC"
              1..728
misc_feature /note="assembly_fragment"
              829..2029
misc_feature /note="assembly_fragment"
              2130..3803
misc_feature /note="assembly_fragment"
              3904..6051
misc_feature /note="assembly_fragment"
              6152..10082
misc_feature /note="assembly_fragment"
              10183..18316
misc_feature /note="assembly_fragment"
              18417..27114
misc_feature /note="assembly_fragment"
              27215..33934
misc_feature /note="assembly_fragment"
              34035..44466
misc_feature /note="assembly_fragment"
              44567..59489
misc_feature /note="assembly_fragment"
              59590..79430
misc_feature /note="assembly_fragment"
              79531..99545
misc_feature /note="assembly_fragment"
              99646..123290
misc_feature /note="assembly_fragment"
              123391..151687
misc_feature /note="assembly_fragment"
              151788..199296
misc_feature /note="assembly_fragment"
              199297..44956
BASE COUNT 53308 a 44956 c 46568 g 53060 t 1404 others
ORIGIN

```

```

Query Match 16.5%; Score 236; DB 2; Length 199296;
Best Local Similarity 100.0%; Pred. No. 3.7e-119;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 806 AGGACACACCGCGGAGACAGACAGCACTGTTAAACGAGAGTTGAGACCTGTGGCGGG 865
      |
Db 124932 AGGACACACCGCGGAGACAGACAGCACTGTTAAACGAGAGTTGAGACCTGTGGCGGG 124931
      |
OY 866 TTTTGAGAGAGCGCAAGGCGCAACTGCTTCAACGACAGTGGCGGAGAGAGAGAGT 925
      |
Db 124992 TTTTGAGAGAGCGCAAGGCGCAACTGCTTCAACGACAGTGGCGGAGAGAGAGAGT 125051
      |
OY 926 TGCACGCGGTCGCGGCGCTCATCGCAGTACGAGAGACCACTTGGAGGGCTCTCAAGC 985
      |
Db 125052 TGCACGCGGTCGCGGCGCTCATCGCAGTACGAGAGACCACTTGGAGGGCTCTCAAGC 125111
      |
OY 986 TGGTGGAGTCCGCGCATCGTCCATGAGAGAGCGCGAGTGGCTCTCACTCCAG 1041
      |
Db 125112 TGGTGGAGTCCGCGCATCGTCCATGAGAGAGCGCGAGTGGCTCTCACTCCAG 125167
      |

```

```

RESULT 5 AC114541/c 49743 bp DNA linear HTG 10-MAR-2002
LOCUS AC114541
DEFINITION Mus musculus clone RP23-8963, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC114541
VERSION AC114541.1 GI:19311061
KEYWORDS HTG; HTGS PHASE0.
SOURCE Mus musculus.
ORGANISM Mus musculus.
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 49743)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
          Unpublished
          2 (bases 1 to 49743)
REFERENCE Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
          Unpublished
AUTHORS

```

# TITLE JOURNAL COMMENT

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,S., Dodge,S., Fairo,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,D., Gadyua,S., Ginde,S., Gord,S., Goyette,M., Graham,D., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamet,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., MacDonald,P., Major,J., Margis,N., Matthews,C., McCarty,M., McEwan,P., McKernan,K., Meldrum,D., Meneus,L., Mihova,T., Mlenka,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunhaph,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thoman,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Triggilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (10-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit,A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L17588  
Center clone name: 89\_G-3

----- NOTE: This record contains 62 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1  
\* 714 813: contig of 713 bp in length  
\* 814 1529: contig of 716 bp in length  
\* 1530 1629: gap of 100 bp  
\* 1630 2320: contig of 691 bp in length  
\* 2321 2420: gap of 100 bp  
\* 2421 3085: contig of 665 bp in length  
\* 3086 3185: gap of 100 bp  
\* 3186 3882: contig of 697 bp in length  
\* 3883 3982: gap of 100 bp  
\* 3983 4708: contig of 726 bp in length  
\* 4709 4808: gap of 100 bp  
\* 4809 5518: contig of 710 bp in length  
\* 5519 5618: gap of 100 bp  
\* 5619 6319: contig of 701 bp in length  
\* 6320 6419: gap of 100 bp  
\* 6420 7128: contig of 709 bp in length  
\* 7129 7228: gap of 100 bp  
\* 7229 7923: contig of 695 bp in length  
\* 7924 8023: gap of 100 bp  
\* 8024 8725: contig of 702 bp in length  
\* 8726 8825: gap of 100 bp  
\* 8826 9527: contig of 702 bp in length  
\* 9528 9627: gap of 100 bp  
\* 9628 10311: contig of 684 bp in length  
\* 10312 10411: gap of 100 bp



10412 11144: contig of 733 bp in length  
11145 11244: gap of 100 bp  
11145 11929: contig of 685 bp in length  
11930 12029: gap of 100 bp  
12030 12728: contig of 700 bp in length  
12730 12829: gap of 100 bp  
12830 13553: contig of 724 bp in length  
13554 13653: gap of 100 bp  
13654 14359: contig of 706 bp in length  
14360 14459: gap of 100 bp  
14460 15161: contig of 702 bp in length  
15162 15261: gap of 100 bp  
15262 15950: contig of 689 bp in length  
15951 16050: gap of 100 bp  
16051 16749: contig of 699 bp in length  
16750 16849: gap of 100 bp  
16850 17542: contig of 693 bp in length  
17543 17642: gap of 100 bp  
17643 18352: contig of 710 bp in length  
18353 18452: gap of 100 bp  
18453 19168: contig of 716 bp in length  
19169 19268: gap of 100 bp  
19269 19893: contig of 715 bp in length  
19894 20083: gap of 100 bp  
20084 20778: contig of 695 bp in length  
20779 20878: gap of 100 bp  
20879 21577: contig of 699 bp in length  
21578 21677: gap of 100 bp  
21678 22380: contig of 703 bp in length  
22381 22480: gap of 100 bp  
22481 23195: contig of 715 bp in length  
23196 23295: gap of 100 bp  
23296 24011: contig of 716 bp in length  
24012 24111: gap of 100 bp  
24112 24824: contig of 713 bp in length  
24825 24924: gap of 100 bp  
24925 25636: contig of 712 bp in length  
25637 25736: gap of 100 bp  
25737 26448: contig of 712 bp in length  
26449 26548: gap of 100 bp  
26549 27255: contig of 707 bp in length  
27256 27355: gap of 100 bp  
27356 28061: contig of 706 bp in length  
28062 28161: gap of 100 bp  
28162 28866: contig of 705 bp in length  
28867 28966: gap of 100 bp  
28967 29664: contig of 698 bp in length  
29665 29764: gap of 100 bp  
29765 30465: contig of 701 bp in length  
30466 30565: gap of 100 bp  
30566 31275: contig of 710 bp in length  
31276 31375: gap of 100 bp  
31376 32084: contig of 709 bp in length  
32085 32184: gap of 100 bp  
32185 32874: contig of 690 bp in length  
32875 32974: gap of 100 bp  
32975 33679: contig of 705 bp in length  
33680 33779: gap of 100 bp  
33780 34485: contig of 706 bp in length  
34486 34585: gap of 100 bp  
34586 35279: contig of 694 bp in length  
35280 35379: gap of 100 bp  
35380 36062: contig of 683 bp in length  
36063 36162: gap of 100 bp  
36163 36856: contig of 694 bp in length  
36857 36956: gap of 100 bp  
36957 37656: contig of 700 bp in length  
37657 37756: gap of 100 bp  
37757 38460: contig of 704 bp in length  
38461 38560: gap of 100 bp  
38561 39273: contig of 713 bp in length  
39274 39373: gap of 100 bp  
39374 40097: contig of 724 bp in length

40098 40197: gap of 100 bp  
40198 40889: contig of 692 bp in length  
40890 40989: gap of 100 bp  
40990 41692: contig of 703 bp in length  
41693 41792: gap of 100 bp  
41793 42505: contig of 713 bp in length  
42506 42605: gap of 100 bp  
42606 43318: contig of 713 bp in length  
43319 43418: gap of 100 bp  
43419 44124: contig of 706 bp in length  
44125 44224: gap of 100 bp  
44225 44923: contig of 699 bp in length  
44924 45023: gap of 100 bp  
45024 45739: contig of 716 bp in length  
45740 45839: gap of 100 bp  
45840 46534: contig of 695 bp in length  
46535 46634: gap of 100 bp  
46635 47333: contig of 699 bp in length  
47334 47433: gap of 100 bp  
47434 48138: contig of 705 bp in length  
48139 48238: gap of 100 bp  
48239 48937: contig of 699 bp in length  
48938 49037: gap of 100 bp  
49038 49743: contig of 706 bp in length.  
location/Qualifiers  
1. 49743  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-8963"  
/clone\_lib="RPCI-23 Female Mouse BAC"  
BASE COUNT 10978 a 10783 c 10620 g 11215 t 6147 others  
ORIGIN  
Query Match 12.1%; Score 173; DB 2; Length 49743;  
Best Local Similarity 100.0%; Pred. No. 3,8e-84;  
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 539 GGGCAGTGCAGCGCCAGGCTGACAGCAGCCTCAGTGAGAGCAGCAGAGAGA 598  
|||||  
DB 3661 GGGCAGTGCAGCGCCAGGCTGACAGCAGCCTCAGTGAGAGCAGCAGAGAGA 3602  
OY 599 TCACATCTACTGCTGAGTGCAGAGGTGCCACCTGCTCTCTGCAAGGTTTGGCG 658  
|||||  
DB 3601 TCACATCTACTGCTGAGTGCAGAGGTGCCACCTGCTCTCTGCAAGGTTTGGCG 3542  
OY 659 CCCACAGGAGCTGTGAGTGGCCCCCTCTGCCACATTTCACAAAGCCAGAAG 711  
|||||  
DB 3541 CCCACAGGAGCTGTGAGTGGCCCCCTCTGCCACATTTCACAAAGCCAGAAG 3489  
RESULT 6  
AC095184 164772 bp DNA linear HTG 11-JUL-2002  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-9D7, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
AC095184 85 unordered pieces.  
AC095184.3 GI:21722570  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 164772)  
REFERENCE  
AUTHORS  
Muzny,D.M., Adams,C., Adio-Oduola,B., All-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barbaria,J., Benton,J., Blimage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brileya,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Doutharte,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Geo,U., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
Homsí,F., Howard,S., Huber,J., Huliyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Kovach,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,  
Lozadó,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhinney,E., Mcleod,M.P., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Nickerson,E., Nwokenkwu,S., Oguh,M., Okwunna,G.,  
Nguyen,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Oragunye,N., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherer,S., Scott,G., Shen,H., Shooshair,N., Sisson,I.,  
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Umanat,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gbbs,R.

Direct Submission  
2 (bases 1 to 164772)  
Worley,K.C.

Direct Submission  
Submitted (16-SEP-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 164772)  
Worley,K.C.

Direct Submission  
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 10, 2002 this sequence version replaced gi:17942274.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GPAM  
Center clone name: CH230-9D7  
----- Summary Statistics  
Sequencing vector: Plasmid  
Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap: version 0.990329  
Consensus quality: 79000 bases at least Q40  
Consensus quality: 85421 bases at least Q30  
Consensus quality: 90756 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see <http://www.hgsc.bcm.tmc.edu/docs/Genbankdraftdata.html>).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 85 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1297: contig of 1297 bp in length

1298 1397: gap of unknown length  
1398 2679: contig of 1282 bp in length  
2680 2779: gap of unknown length  
2780 3802: contig of 1023 bp in length  
3803 3902: gap of unknown length  
3903 4951: contig of 1049 bp in length  
4952 5051: gap of unknown length  
5052 6239: contig of 1248 bp in length  
6300 6399: gap of unknown length  
6400 7431: contig of 1032 bp in length  
7432 7531: gap of unknown length  
7532 8548: contig of 1017 bp in length  
8549 9776: contig of 1128 bp in length  
9777 9876: gap of unknown length  
9877 10877: contig of 1001 bp in length  
10878 10977: gap of unknown length  
10978 12067: contig of 1090 bp in length  
12068 12167: gap of unknown length  
12168 13239: contig of 1072 bp in length  
13240 13339: gap of unknown length  
13340 14779: contig of 1439 bp in length  
14779 14878: gap of unknown length  
14879 16415: contig of 1537 bp in length  
16416 17592: gap of unknown length  
17592 17692: contig of 1077 bp in length  
17693 18970: gap of unknown length  
18970 19070: contig of 1278 bp in length  
19071 20403: gap of unknown length  
20404 20503: contig of 1333 bp in length  
20504 21651: gap of unknown length  
21652 21751: contig of 1148 bp in length  
21752 22887: gap of unknown length  
22887 22987: contig of 1136 bp in length  
22988 22987: gap of unknown length  
22988 24446: contig of 1459 bp in length  
24447 24546: gap of unknown length  
24547 25565: contig of 1019 bp in length  
25566 27036: gap of unknown length  
27037 27136: contig of 1371 bp in length  
27137 28478: gap of unknown length  
28479 28578: contig of 1342 bp in length  
28579 29664: gap of unknown length  
29665 29764: contig of 1086 bp in length  
29765 31068: gap of unknown length  
31069 31168: contig of 1304 bp in length  
31169 32406: gap of unknown length  
32407 32506: contig of 1238 bp in length  
32507 34145: gap of unknown length  
34146 34245: contig of 1639 bp in length  
34246 35811: gap of unknown length  
35812 35911: contig of 1566 bp in length  
35912 37018: gap of unknown length  
37019 37118: contig of 1107 bp in length  
37119 38624: gap of unknown length  
38625 38724: contig of 1506 bp in length  
38725 40409: gap of unknown length  
40410 40509: contig of 1685 bp in length  
40510 42002: gap of unknown length  
42003 42102: contig of 1493 bp in length  
42103 43258: gap of unknown length  
43159 43258: contig of 1056 bp in length  
43259 44262: gap of unknown length  
44263 44362: contig of 1004 bp in length  
44363 45368: gap of unknown length  
45369 45668: contig of 1206 bp in length  
45669 47129: gap of unknown length  
47130 47229: contig of 1461 bp in length  
47230 48940: gap of unknown length  
48941 49040: contig of 1711 bp in length  
49041 50571: gap of unknown length  
50572 50671: contig of 1531 bp in length  
50671: gap of unknown length

\* 50672 52207: contig of 1536 bp in length  
\* 52208 52307: gap of unknown length  
\* 52308 52349: contig of 1142 bp in length  
\* 53450 53459: gap of unknown length  
\* 53450 54869: contig of 1320 bp in length  
\* 54870 54969: gap of unknown length  
\* 54970 56026: contig of 1057 bp in length  
\* 56027 56126: gap of unknown length  
\* 56127 58611: contig of 2485 bp in length  
\* 58612 58712: gap of unknown length  
\* 58713 60042: contig of 1331 bp in length  
\* 60043 60142: gap of unknown length  
\* 60143 61451: contig of 1309 bp in length  
\* 61452 61551: gap of unknown length  
\* 61552 63614: contig of 2063 bp in length  
\* 63615 63714: gap of unknown length  
\* 63715 65157: contig of 1443 bp in length  
\* 65158 65257: gap of unknown length  
\* 65258 66957: contig of 1700 bp in length  
\* 66958 67057: gap of unknown length  
\* 67058 69461: contig of 2404 bp in length  
\* 69462 70953: gap of unknown length  
\* 70954 71053: gap of unknown length  
\* 71054 72992: contig of 1939 bp in length  
\* 72993 73092: gap of unknown length  
\* 73093 75634: contig of 2542 bp in length  
\* 75635 75734: gap of unknown length  
\* 75735 77844: contig of 2110 bp in length  
\* 77845 77944: gap of unknown length  
\* 77945 80323: contig of 2379 bp in length  
\* 80324 80423: gap of unknown length  
\* 80424 81947: contig of 1524 bp in length

Query Match 4.0%; Score 57; DB 2; Length 164772;  
Best Local Similarity 100.0%; Pred. No. 1.2e-19;  
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 655 GGGCCCAAGAGACTGTGAGTGGCCCTCGCCACCATTTACAAAGCGCAGAG 711  
Db 162273 GGGCCCAAGAGACTGTGAGTGGCCCTCGCCACCATTTACAAAGCGCAGAG 162329

RESULT 7  
LOCUS HSA291714 1329 bp mRNA linear PRI 06-JUL-2001  
DEFINITION Homo sapiens mRNA for RNF30 gene for ring finger protein 30.  
ACCESSION AJ291714.1 GI:13160387  
VERSION MURF-3 gene; muscle specific ring finger protein.  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 1329)  
Centner,T., Yano,J., Kimura,E., McElhinny,A.S., Pelin,K.,  
Wilt,C.C., Bang,M.L., Trombly,K., Granzier,H., Gregorio,C.C.,  
Sornichsen,H. and Labelle,S.  
Identification of muscle specific ring finger proteins as potential  
regulators of the titin kinase domain  
J. Mol. Biol. 306 (4), 717-726 (2001)  
JOURNAL MEDLINE 21140140  
PUBMED 11243782  
REFERENCE 2 (bases 1 to 1329)  
Centner,T.  
Direct Submission  
Submitted (12-FEB-2001) Centner T., Structure and Biocomputing,  
EMBL-Heidelberg, Meyerhofstrasse 1, Heidelberg 69112, Germany  
Revised by author 22-FEB-2001  
REMARK FEATURES  
Location/Qualifiers  
1..1329  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

misc-feature 1..3  
/note="putative alternative translation initiation codon"  
/evidence=not\_experimental  
gene 49..1203  
/gene="RNF30"  
CDS 49..1203  
/gene="RNF30"  
/note="alternative"  
/product="ring finger protein 30"  
/protein\_id="CAC32842.1"  
/db\_xref="GI:13160389"

CDS  
/translation="MDNLEKQLICPICLEMFSKPVVILPCQHNLCRCANDVPOASNP  
LWQSRSTVSSGRCPCREHYVLDHRYVGLQNLIVENTIDYKQSSRPLHS  
KEOHLMCEHHEEKINIVCLSCPEVSLCKYGAHKDEVAFLPTIYRKQKODLT  
LPLRLCSGNTTYCSIDLPSSDPPLASQNKIIDSSELDGJAMLVAGNDROAVYI  
TOMEVFCOTTEPNSRRKOLLQRFESICVLEERKGLQALAREQEKIQRVGLI  
RQGDHLEASSKIVESAIOSMEROMALYIQAKELINKGAMSKVELAGPPEGYES  
MEQTVRVEHVAEMLRITDFPGASGEHEEVADPGEESAGPEERPDGP"  
join(49..515,642..1203)  
/gene="RNF30"  
/note="short isoform"  
/codon\_start=1  
/product="ring finger protein 30"  
/protein\_id="CAC32841.1"  
/db\_xref="GI:13160388"

BASE COUNT 310 a 375 c 414 g 230 t

Query Match 2.5%; Score 36; DB 9; Length 1329;  
Best Local Similarity 100.0%; Pred. No. 5.1e-08;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1099 CGGCCGAGCCAGGCTATGAGAGCATGAGCAATTC 1134  
Db 1027 CGGCCGAGCCAGGCTATGAGAGCATGAGCAATTC 1062

RESULT 8  
LOCUS AC013413 118847 bp DNA linear PRI 21-FEB-2002  
DEFINITION Homo sapiens BAC clone RP11-538J11 from 2, complete sequence.  
ACCESSION AC013413  
VERSION AC013413.6 GI:18158397  
KEYWORDS HMG.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 118847)  
Sulston,J.E. and Waterston,R.  
Toward a complete human genome sequence  
Genome Res. 8 (11), 1097-1108 (1998)  
JOURNAL MEDLINE 99063792  
PUBMED 9847074  
REFERENCE 2 (bases 1 to 118847)  
Swearengen,S. and Kozlowski,A.  
The sequence of Homo sapiens BAC clone RP11-538J11  
Unpublished (2001)  
REFERENCE 3 (bases 1 to 118847)  
Waterston,R.H.  
Direct Submission  
Submitted (09-NOV-1999) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
4 (bases 1 to 118847)

AUTHORS Waterston, R. H.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 5 (bases 1 to 118847)  
 WATERSTON, R.  
 DIRECT SUBMISSION  
 Submitted (21-FEB-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Jan 10, 2002 this sequence version replaced g1:16418272.  
 -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc  
 Contact: saplens@wustl.wustl.edu  
 -----  
 Summary Statistics  
 Center project name: H\_NH0538J11  
 -----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:  
 The RPCR-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenoe, M., Catalanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org  
 VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:  
 The clone sequenced to the left is RP11-195B17, 2000 bp overlap. Actual start of this clone is at base position 80914 of RP11-195B17; actual end is at base position 118847 of RP11-538J11. RP11-538J11 contains a transposon in the growth of the clone which has been omitted from the finished sequence. Data from AC025662 was used to finish this clone, AC013413.

#### FEATURES

source  
 1. 118847  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="2"  
 /map="2"  
 /clone="RP11-538J11"  
 /clone\_lib="RPCR-11"  
 1. 1159  
 /note="match to EST BE789511 (NID:g10210709)"  
 1. 1151  
 /note="match to EST A1978781 (NID:g5803811) wr59f04.x1"  
 133. 154  
 /note="match to EST BG256909 (NID:g12766725)"  
 393. 626  
 /note="similar to Mus musculus EST BB659249"

misc\_feature (NID:g16493072)"  
 404. 626  
 /note="similar to Mus musculus EST B1661519 (NID:g1557575)"  
 691. 700  
 /note="match to EST A1905917 (NID:g6496304)"  
 870. 1009  
 /note="similar to Mus musculus EST BB659249 (NID:g16493072)"  
 870. 1009  
 /note="similar to Mus musculus EST B1661519 (NID:g1557575)"  
 1991. 2327  
 /note="match to EST AA758920 (NID:g2806783) ah77b09.sl"  
 2336. 2554  
 /rpt\_family="Alu"  
 2551. 2600  
 /rpt\_family="T-rich"  
 2576. 2880  
 /rpt\_family="Alu"  
 2971. 3156  
 /rpt\_family="L2"  
 3359. 3667  
 /rpt\_family="Alu"  
 3669. 3979  
 /rpt\_family="Alu"  
 3951. 3979  
 /rpt\_family="AT-rich"  
 4211. 4340  
 /note="similar to Mus musculus EST BB659249 (NID:g16493072)"  
 4211. 4340  
 /note="similar to Mus musculus EST B1661519 (NID:g1557575)"  
 4529. 5049  
 /rpt\_family="L1"  
 4747. 4773  
 /rpt\_family="(T)n"  
 5187. 5329  
 /note="similar to Mus musculus EST B1661519 (NID:g1557575)"  
 5187. 5309  
 /note="similar to Mus musculus EST BB659249 (NID:g16493072)"  
 5512. 5651  
 /note="similar to Mus musculus EST B1661519 (NID:g1557575)"  
 5528. 5657  
 /note="similar to Homo sapiens EST BF761009 (NID:g12108987)"  
 5552. 5653  
 /note="match to EST BE242145 (NID:g9093794)"  
 5552. 5653  
 /note="match to EST BM152742 (NID:g17177267)"  
 5553. 5653  
 /note="match to EST BM152742 (NID:g17177267)"  
 5553. 5653  
 /note="similar to Mus musculus EST BF532585 (NID:g11619948)"  
 5859. 6030  
 /note="match to EST BM152742 (NID:g17177267)"  
 5859. 6030  
 /note="similar to Mus musculus EST BF532585 (NID:g11619948)"  
 5859. 5931  
 /note="match to EST BE242145 (NID:g9093794)"  
 5867. 5912  
 /note="similar to Homo sapiens EST BF761009 (NID:g12108987)"  
 5942. 6030  
 /note="match to EST AT244159 (NID:g3839556) qj80e03.x1"  
 6145. 6450  
 /rpt\_family="Alu"  
 6428. 6450  
 /note="match to EST BE789511 (NID:g10210709)"

misc\_feature 6556..6741  
/note="match to EST AI244159 (NID:q3839556) qj80e03.x1"  
misc\_feature 6556..6741  
/note="similar to Mus musculus EST BF532585  
(NID:q11619948)"  
misc\_feature 6556..6692  
/note="match to EST BM152742 (NID:q17177267)"  
misc\_feature 6910..7022  
/note="match to EST AI244159 (NID:q3839556) qj80e03.x1"  
misc\_feature 6910..7022  
/note="similar to Mus musculus EST BF532585  
(NID:q11619948)"  
repeat\_region 7252..7332  
/rpt\_family="MIR"  
misc\_feature 7338..7483  
/note="similar to Mus musculus EST BF532585  
(NID:q11619948)"  
misc\_feature 7338..7420  
/note="match to EST AI244159 (NID:q3839556) qj80e03.x1"  
misc\_feature 7738..7817  
/note="similar to Mus musculus EST BF532585  
(NID:q11619948)"  
repeat\_region 8289..8603  
/rpt\_family="Alu"  
repeat\_region 8292..8316  
/rpt\_family="(?)n"  
misc\_feature 8866..8919  
/note="similar to Mus musculus EST B1647507"

Query Match 2.5%; Score 36; DB 9; Length 118847;  
Best Local Similarity 100.0%; Pred. No. 5.7e-08;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1099 CGCCGAGCAGGCTATGAGCATGAGCAATTC 1134  
DB 89240 CGCCGAGCAGGCTATGAGCATGAGCAATTC 89275

RESULT 9  
LOCUS AX060632 1500 bp DNA linear PAT 22-JAN-2001  
DEFINITION Sequence 54 from Patent WO0078954.  
ACCESSION AX060632  
VERSION AX060632.1 GI:12406053  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1500)  
AUTHORS Lal,P., Yue,H., Tang,Y.T., Baughn,M.R., Azimzai,Y. and Tran,B.  
TITLE Human transcriptional regulator proteins  
JOURNAL Patent: WO 0078954-A 54 28-DEC-2000;  
Incyte Genomics, Inc. (US)

FEATURES  
Source location/Qualifiers  
1..1500  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 401 a 378 c 439 g 282 t  
ORIGIN

Query Match 2.4%; Score 35; DB 6; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 1.8e-07;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 685 CAGCCACATTTACAAAGCCAGAGAGTACGT 719  
DB 777 CTGCCACCATTTACAAAGCCAGAGAGTACGT 811

RESULT 10  
LOCUS AC025662/AC025662/c 123280 bp. DNA linear HTG 13-APR-2000

DEFINITION Homo sapiens clone RP11-352122, LOW-PASS SEQUENCE SAMPLING.  
AC025662  
VERSION AC025662.2 GI:7547186  
KEYWORDS HTG; PHASEO.  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 123280)  
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
TITLE Homo sapiens, clone RP11-352122  
JOURNAL Unpublished  
2 (bases 1 to 123280)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bede,F.,  
Boguslavsky,L., Bouckgeater,B., Brown,A., Burkett,G.,  
Campoliano,A., Castle,A., Chappel,Y., Colangelo,M., Collins,S.,  
Collins,A., Cooke,P., Dearlano,K., Dewar,K., Diaz,J.S.,  
Dodg,S., Domno,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Glnde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,  
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
Klein,J., Labocque,K., Lamazares,R., Landers,T., Lechoczy,J.,  
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
McCarthy,M., McEwan,P., McGuire,A., McKernan,K., McNeeters,R.,  
McElroy,J., Menes,L., Mihova,T., Miranda,C., Mianga,V., Morrow,J.,  
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,C., Plerre,N.,  
Pisani,C., Pollard,Y., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Stange-Thoman,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Tejeda,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE JOURNAL  
COMMENT Submitted (12-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 13, 2000 this sequence version replaced gi:7230272.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: 352\_1-22  
Center clone name: 352\_1-22

\* NOTE: This record contains 146 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
\* 1  
\* 749 848: gap of 100 bp  
\* 849 1619: contig of 771 bp in length  
\* 1620 1719: gap of 100 bp  
\* 1720 2490: contig of 771 bp in length  
\* 2491 2590: gap of 100 bp  
\* 2591 3378: contig of 788 bp in length  
\* 3379 3478: gap of 100 bp  
\* 3479 4239: contig of 761 bp in length  
\* 4240 4339: gap of 100 bp  
\* 4340 5113: contig of 774 bp in length  
\* 5114 5213: gap of 100 bp

\* 5214 5994: contig of 781 bp in length  
\* 5995 6094: gap of 100 bp  
\* 6095 6873: contig of 779 bp in length  
\* 6874 6973: gap of 100 bp  
\* 6974 7732: contig of 759 bp in length  
\* 7733 7832: gap of 100 bp  
\* 7833 8604: contig of 772 bp in length  
\* 8605 8704: gap of 100 bp  
\* 8705 9478: contig of 774 bp in length  
\* 9479 9578: gap of 100 bp  
\* 9579 10356: contig of 778 bp in length  
\* 10357 10456: gap of 100 bp  
\* 10457 11243: contig of 787 bp in length  
\* 11244 11343: gap of 100 bp  
\* 11344 12125: contig of 782 bp in length  
\* 12126 12225: gap of 100 bp  
\* 12226 13028: contig of 803 bp in length  
\* 13029 13128: gap of 100 bp  
\* 13129 13891: contig of 763 bp in length  
\* 13892 13991: gap of 100 bp  
\* 13992 14763: contig of 772 bp in length  
\* 14764 14863: gap of 100 bp  
\* 14864 15625: contig of 762 bp in length  
\* 15626 15725: gap of 100 bp  
\* 15726 16498: contig of 773 bp in length  
\* 16499 16598: gap of 100 bp  
\* 16599 17376: contig of 778 bp in length  
\* 17377 17476: gap of 100 bp  
\* 17477 18246: contig of 770 bp in length  
\* 18247 18346: gap of 100 bp  
\* 18347 19137: contig of 791 bp in length  
\* 19138 19237: gap of 100 bp  
\* 19238 20026: contig of 789 bp in length  
\* 20027 20126: gap of 100 bp  
\* 20127 20909: contig of 783 bp in length  
\* 20910 21009: gap of 100 bp  
\* 21010 21787: contig of 778 bp in length  
\* 21788 21887: gap of 100 bp  
\* 21888 22671: contig of 784 bp in length  
\* 22672 22771: gap of 100 bp  
\* 22772 23556: contig of 785 bp in length  
\* 23557 23656: gap of 100 bp  
\* 23657 24446: contig of 790 bp in length  
\* 24447 24546: gap of 100 bp  
\* 24547 25324: contig of 778 bp in length  
\* 25325 25424: gap of 100 bp  
\* 25425 26211: contig of 787 bp in length  
\* 26212 26311: gap of 100 bp  
\* 26312 27074: contig of 763 bp in length  
\* 27075 27174: gap of 100 bp  
\* 27175 27960: contig of 786 bp in length  
\* 27961 28060: gap of 100 bp  
\* 28061 28825: contig of 765 bp in length  
\* 28826 28925: gap of 100 bp  
\* 28926 29722: contig of 797 bp in length  
\* 29723 29822: gap of 100 bp  
\* 29823 30598: contig of 776 bp in length  
\* 30599 30698: gap of 100 bp  
\* 30699 31472: contig of 774 bp in length  
\* 31473 31572: gap of 100 bp  
\* 31573 32358: contig of 786 bp in length  
\* 32359 32458: gap of 100 bp  
\* 32459 33227: contig of 769 bp in length  
\* 33228 33327: gap of 100 bp  
\* 33328 34098: contig of 771 bp in length  
\* 34099 34198: gap of 100 bp  
\* 34199 34978: contig of 780 bp in length  
\* 34979 35078: gap of 100 bp  
\* 35079 35867: contig of 789 bp in length  
\* 35868 35967: gap of 100 bp  
\* 35968 36782: contig of 815 bp in length  
\* 36783 36882: gap of 100 bp  
\* 36883 37630: contig of 748 bp in length

\* 37631 37730: gap of 100 bp  
\* 37731 38499: contig of 769 bp in length  
\* 38500 38599: gap of 100 bp  
\* 38600 39388: contig of 789 bp in length  
\* 39389 39488: gap of 100 bp  
\* 39489 40255: contig of 767 bp in length  
\* 40256 40355: gap of 100 bp  
\* 40356 41129: contig of 774 bp in length  
\* 41130 41229: gap of 100 bp  
\* 41230 41998: contig of 769 bp in length  
\* 41999 42098: gap of 100 bp  
\* 42099 42868: contig of 770 bp in length  
\* 42869 42968: gap of 100 bp  
\* 42969 43746: contig of 778 bp in length  
\* 43747 43846: gap of 100 bp  
\* 43847 44632: contig of 786 bp in length  
\* 44633 44732: gap of 100 bp  
\* 44733 45522: contig of 790 bp in length  
\* 45523 45622: gap of 100 bp  
\* 45623 46403: contig of 781 bp in length  
\* 46404 46503: gap of 100 bp  
\* 46504 47312: contig of 809 bp in length  
\* 47313 47412: gap of 100 bp  
\* 47413 48189: contig of 777 bp in length  
\* 48190 48289: gap of 100 bp  
\* 48290 49052: contig of 763 bp in length  
\* 49053 49152: gap of 100 bp  
\* 49153 49929: contig of 777 bp in length  
\* 49930 50029: gap of 100 bp  
\* 50030 50795: contig of 766 bp in length  
\* 50796 50895: gap of 100 bp  
\* 50896 51681: contig of 786 bp in length  
\* 51682 51781: gap of 100 bp  
\* 51782 52563: contig of 782 bp in length  
\* 52564 52663: gap of 100 bp  
\* 52664 53443: contig of 780 bp in length  
\* 53444 53543: gap of 100 bp  
\* 53544 54331: contig of 788 bp in length  
\* 54332 54431: gap of 100 bp  
\* 54432 55216: contig of 785 bp in length  
\* 55217 55316: gap of 100 bp  
\* 55317 56115: contig of 799 bp in length  
\* 56116 56215: gap of 100 bp  
\* 56216 57005: contig of 790 bp in length  
\* 57006 57105: gap of 100 bp  
\* 57106 57874: contig of 769 bp in length  
\* 57875 57974: gap of 100 bp  
\* 57975 58750: contig of 776 bp in length  
\* 58751 58850: gap of 100 bp  
\* 58851 59618: contig of 768 bp in length  
\* 59619 59718: gap of 100 bp  
\* 59719 60490: contig of 772 bp in length  
\* 60491 60590: gap of 100 bp

Query Match 2.0%; Score 29; DB 2; Length 123280;  
Best Local Similarity 100.0%; Pred. No. 0.00044;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 622 GAGTGGCCACGTCGTCGTCGACAGGT 650  
Db 32873 GAGTGGCCACGTCGTCGTCGACAGGT 32845

RESULT 11  
AX418852  
LOCUS AX418852  
DEFINITION Sequence 5 from Patent WO0206318.  
ACCESSION AX418852  
VERSION AX418852.1 GI:21523716  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 AUTHORS Olson, E.N. and Spencer, J.A.  
 TITLE Methods and compositions for stabilizing microtubules and intermediate filaments in striated muscle cells  
 JOURNAL Patent: WO 0206318-A 5 24-JAN-2002;  
 BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)  
 FEATURES Location/Qualifiers  
 source 1.1597  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 299.1330  
 /note="unnamed protein product"  
 /protein\_id="CAD35457.1"  
 /db\_xref="GI:21523717"  
 /translation="MDYKSSLIIPDGNMENELOICPICLEMFKPVILLPCQNLIC  
 RKANDIFFOANPYWTKNRGSGVSMGSRFRPSCREHVIIMDRHGYIGLQRLNVNII  
 DIYKQECSSRPLOKSHPMCKEHEDEKINITYCLCEVPYSLCKVGAHQACEVAPLO  
 SITQOGKTELSTNCISMLVAGNDRVQITISQLEDSRVKRENSHVKELSGKFDPLA  
 ILDEKKSSELLQRIQEOEKEIGFLEALIIQYREOLEKSTKIVETATISQLEDEGATFL  
 SSKOLIKSIYEASKGCGOLGTEGCFEMDMYFTLDLHIAELRAIDFGTGKGCVCVC  
 LFEERORSS"  
 BASE COUNT 405 a 404 c 438 g 350 t  
 ORIGIN

Query Match 2.0%; Score 28; DB 6; Length 1597;  
 Best Local Similarity 100.0%; Pred. No. 0.0014;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 505 AACATCATGTGACATCTACACAGCAGACT 532  
 ||||||||||||||||||||||||||||  
 Db 596 AACATCATGTGACATCTACACAGCAGACT 623

RESULT 12  
 LOCUS AL627314 196461 bp DNA linear ROD 09-JUL-2002  
 DEFINITION Mouse DNA sequence from clone RP23-354H24 on chromosome 4, complete  
 sequence.  
 ACCESSION AL627314  
 VERSION AL627314.6 GI:21727344  
 KEYWORDS HTG.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 196461)  
 BROWN, A.  
 Direct Submission  
 Submitted (09-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquery@sanger.ac.uk Clone requests: clonesrequest@sanger.ac.uk  
 On Jul 10, 2002 this sequence version replaced g1:1815285.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP  
 database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep RP23-354H24 is  
 from the RPI-23 mouse PAC Library

constructed by the group of Pieter de Jong.  
 For further details see <http://www.chori.org/jacpac/home.htm>  
 VECTOR: pBACE3.6  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: humquery@sanger.ac.uk  
 -----

FEATURES  
 source 1.196461  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="4"  
 /clone="RP23-354H24"  
 /clone\_lib="RPI-23"  
 BASE COUNT 51927 a 47767 c 46054 g 50713 t  
 ORIGIN

Query Match 2.0%; Score 28; DB 10; Length 196461;  
 Best Local Similarity 100.0%; Pred. No. 0.0016;  
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 505 AACATCATGTGACATCTACACAGCAGACT 532  
 ||||||||||||||||||||||||||||  
 Db 188387 AACATCATGTGACATCTACACAGCAGACT 188414

RESULT 13  
 LOCUS AB047601 1746 bp mRNA linear PRI 11-OCT-2001  
 DEFINITION Macaca fascicularis brain cDNA, clone:Onpa-10466.  
 ACCESSION AB047601  
 VERSION AB047601.1 GI:9929936  
 KEYWORDS fit (full insert sequence).  
 SOURCE Macaca fascicularis adult male brain parietal lobe cDNA to mRNA,  
 clone\_lib:macaque brain cDNA library Onpa clone:Onpa-10466.  
 ORGANISM Macaca fascicularis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 Cercopithecinae; Macaca.  
 1  
 Osada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirata, M.,  
 Suto, Y., Hirai, M., Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.  
 Assignment of 118 novel cDNAs of cynomolgus monkey brain to human  
 chromosomes  
 Gene 275 (1), 31-37 (2001)  
 2 (bases 1 to 1746)  
 21458551  
 Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.  
 Direct Submission  
 Submitted (23-AUG-2000) Katsuyuki Hashimoto, National Institute of  
 Infectious Diseases, Division of Genetic Resources; 23-1, Toyama  
 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan  
 (E-mail: khashimn@nig.igp.urfu: <http://www.nih.go.jp/yoken/genebank/>,  
 Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)  
 Lab host: TOP10  
 Vector: pME18S-Fl3 (ACC. NO. AB009864)  
 R. Site1: DraIII (CACTGTGTG)  
 R. Site2: DraIII (CACCATATG)  
 Description: 1st strand cDNA was primed with an oligo(dT) primer  
 [ATGCGCCCTTTTCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized  
 using specific 5' and 3' primers and amplified by PCR. The PCR  
 product was digested with SfiI and size selection was performed to  
 exclude fragments <1.5kb. The SfiI-digested PCR product was cloned  
 into distinct DraIII sites of pME18S-Fl3. XhoI sites just outside  
 the DraIII sites can be used to isolate the cDNA insert. Libraries  
 were constructed by Sugano et al. (University of Tokyo, Institute of  
 Medical Science). Custom primer used for sequencing  
 (5' end primer [CTTCTGCTCTAAAGCTGCG];  
 3' end primer [CGACTGCACTCGAGCA]).  
 Location/Qualifiers  
 1.1746

CDS

```

/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="QnpA-10466"
/sex="male"
/tissue_type="brain parietal lobe"
/clone_lib="macaque brain cDNA library QnpA"
/dev_stage="adult"
567..1244
/codon_start=1
/product="hypothetical protein"
/protein_id="BAB12125.1"
/db_xref="GI:9929937"
/translation="MSASLNYSKFSKEQOTMDNLEKOLICPICLEMTKPVVILPCOH
NICKKASDIFQASNPVLPTRGGTMASSGGRPCSRHEVYLDHGVGLQKRNLYAE
NIDITYKQESTIRPEKSDQPCGEHEBERINITYCLNCEVPTCSLCYFGAHKDCQVAP
LTHVFOROKSELSIDGIALIVGSNDRVGVISOLEDPTCKTIECCRKOKOELCEKFDYL
YGSAGSQ"

```

BASE COUNT 523 a 396 c 387 g 440 t

Query Match 1.9%; Score 27; DB 9; Length 1746;  
Best Local Similarity 100.0%; Pred. No. 0.0052;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 508 ATCATGTGACATCTACAGCAGAGTCC 534  
|||||  
Db 876 ATCATGTGACATCTACAGCAGAGTCC 902

RESULT 14  
LOCUS BC007750 1750 bp mRNA linear PRI 12-JUN-2001  
DEFINITION Homo sapiens, clone MGC:12836 IMAGE:4110783, mRNA, complete cds.  
ACCESSION BC007750.1 GI:14043531  
KEYWORDS MGC.  
SOURCE Homo sapiens.  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1750)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgapbs-rt@mail.nih.gov](mailto:cgapbs-rt@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nigrl.nih.gov](mailto:nisc_mgc@nigrl.nih.gov)  
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lin, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McQuinn, J., Pearson, R., Snyder, B., Stantip, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurugan, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IMAGE Plate: 17 Row: d Column: 4  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.  
Location/Qualifiers

FEATURES

source

1..1750  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:12836 IMAGE:4110783"  
/tissue\_type="Muscle, Rhabdomyosarcoma"  
/clone\_lib="NIH-MGC-17"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
190..1500  
/codon\_start=1  
/product="Unknown (protein for MGC:12836)"  
/protein\_id="AA07750.1"  
/db\_xref="GI:14043532"  
/translation="MDNLEKOLICPICLEMTKPVVILPCOHNLCRKASDIFQASNP  
YLPTRGGTMASSGGRPCSRHEVYLDHGVGLQKRNLYAEINIDITYKQESTIRPEK  
SDQPCGEHEBERINITYCLNCEVPTCSLCYFGAHKDCQVAPLTHVFOROKSELSIDG  
IALIVGSNDRVGVISOLEDPTCKTIECCRKOKOELCEKFDYL  
YGSAGSQ"  
/note="Signal peptide: 1-1750"

CDS

BASE COUNT 556 a 389 c 433 g 372 t

Query Match 1.9%; Score 27; DB 9; Length 1750;  
Best Local Similarity 100.0%; Pred. No. 0.0052;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 508 ATCATGTGACATCTACAGCAGAGTCC 534  
|||||  
Db 451 ATCATGTGACATCTACAGCAGAGTCC 477

RESULT 15  
LOCUS HSA243488 1810 bp mRNA linear PRI 29-JUN-2001  
DEFINITION Homo sapiens titlin zinc-finger anchoring protein, 50kDa isoform.  
ACCESSION AJ243488.1 GI:14588845  
VERSION AJ243488.1 GI:14588845  
KEYWORDS alternative splicing; signal transduction; titlin zinc-finger anchoring protein; titlan.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1810)  
AUTHORS Kelly, R., Neubauer, G. and Gautel, M.  
TITLE A novel RING finger protein associated with titlin kinase  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1810)  
AUTHORS Gautel, M.S.  
TITLE Direct Submission  
JOURNAL Submitted (29-JUN-1999) Gautel, M.S., Physiologische Biochemie, Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse 11, Dortmund, 44227, GERMANY  
Location/Qualifiers

FEATURES

1..1810  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_type="cardiomyocyte"  
/tissue\_type="cardiac muscle"  
227..1585  
/function="putative role in signal transduction"  
/note="50kDa isoform"  
/evidence="experimental"  
/product="titlin zinc-finger anchoring protein"  
/protein\_id="CAC43019.1"  
/db\_xref="GI:14588846"  
/translation="MSASLNYSKFSKEQOTMDNLEKOLICPICLEMTKPVVILPCOH  
NICKKASDIFQASNPVLPTRGGTMASSGGRPCSRHEVYLDHGVGLQKRNLYAE  
NIDITYKQESTIRPEKSDQPCGEHEBERINITYCLNCEVPTCSLCYFGAHKDCQVAP  
LTHVFOROKSELSIDGIALIVGSNDRVGVISOLEDPTCKTIECCRKOKOELCEKFDYL  
YGSAGSQ"



YGLIEKRNEMTOVITRTOEELKLEHVALIRKYSDDLLENVSKLVESGLOFMDPEEMAV  
FLONAKTLKIKISEASKAFOMEKIEHGYENNHNFTVINRBEKIIREIDFYREDEDE  
EEEGGEGKEEVEGEAVEEVENVOTEPEDNEPEKASELSOVELOAAAGALP  
VSSPEPPALPPADAPVTOIGFEAPPLQGAAPASGADSEPARHIFSTWLNLSL  
NE"

BASE COUNT  
ORIGIN

557 a 420 c 448 g 385 t

Query Match 1.9%: Score 27; DB 9; Length 1810;  
Best Local Similarity 100.0%; Pred. No. 0.0052;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 508 ATCATTCACATCTACAGCAGAGTCC 534  
|||||  
536 ATCATTCACATCTACAGCAGAGTCC 562

RESULT 16  
AK091728

LOCUS 1925 bp mRNA linear PRI 15-JUL-2002  
DEFINITION Homo sapiens CDNA FLJ34409 fis, clone HEART2001931, moderately  
similar to Mus musculus RING-finger protein MURF mRNA.  
AK091728  
AK091728.1 GI:21750167  
oligo capping: fis (full insert sequence).  
KEYWORDS Homo sapiens heart CDNA to mRNA, clone\_lib:HEART2  
SOURCE clone:HEART2001931.  
Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

ORGANISM

REFERENCE  
AUTHORS

1 Nishi, T., Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,  
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,  
Yamamoto, J., Isono, Y., Kawai, H., Saito, K., Nishikawa, T.,  
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,  
Kikuchi, H., Kanda, K., Magatsu, M., Murakawa, K., Kanehori, K.,  
Takahashi, Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,  
Sugano, S., Nagahara, K., Masuno, Y., Nagai, K. and Isogai, T.  
NDO human CDNA sequencing project  
Unpublished  
2 (bases 1 to 1925)  
Isogai, T. and Yamamoto, J.  
Direct Submission  
Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7  
Kazuma-Kamatairi, Kisarazu, Chiba 292-0812, Japan  
E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NDO human CDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; CDNA full insert sequencing:  
Research Association for Biotechnology (RAB); CDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: HRI and  
RAB; annotation: HRI and RAB.  
Location/Qualifiers  
1. 1925  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="HEART2001931"  
/tissue\_type="heart"  
/clone\_lib="HEART2"  
/note="cloning vector: PME18SFL3"

COMMENT  
JOURNAL

REFERENCE  
AUTHORS

TITLE

ORIGIN

BASE COUNT

ORIGIN

579 a 441 c 498 g 407 t

Query Match 1.9%: Score 27; DB 9; Length 1925;  
Best Local Similarity 100.0%; Pred. No. 0.0052;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 508 ATCATTCACATCTACAGCAGAGTCC 534  
|||||  
536 ATCATTCACATCTACAGCAGAGTCC 382

RESULT 17  
HSA243489

2098 bp mRNA linear PRI 29-JUN-2001  
HSA243489  
Homo sapiens titin zinc-finger anchoring protein, 60kDa isoform.

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

aj243489.1 GI:14588847  
alternative splicing: signal transduction; titin zinc-finger  
anchoring protein; titian.  
human.  
Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

SOURCE

human.  
Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 2098)  
Kelly, R., Neubauer, G. and Gautel, M.  
A novel RING finger protein associated with titin kinase  
Unpublished  
2 (bases 1 to 2098)  
Gautel, M.S.  
Direct Submission  
Submitted (29-JUN-1999) Gautel M.S., Physiologische Biochemie,  
Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse  
11, Dortmund, 44227, GERMANY  
Location/Qualifiers  
1. 2098  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_type="cardiomyocyte"  
/tissue\_type="cardiac muscle"  
227. 1873  
/function="putative role in signal transduction"  
/note="60kDa isoform"  
/codon\_start=1  
/evidence="experimental"  
/product="titin zinc-finger anchoring protein"  
/protein\_id="CAC43020.1"  
/db\_xref="GI:14588848"  
/translation="MSASLVNYSFSPKRODMNLEKOLICPLEMFTKPVIIIPCOH  
NIDITKQESTREPKKSDQPCHEHEERINITOLNCEVPCISCKYFGAKHDCQVAP  
LTHVFORKSELSDDGIALIVGSNDVQGVISQEDTCTIECCRKQOECEKEDVL  
YGLIEKRNEMTOVITRTOEELKLEHVALIRKYSDDLLENVSKLVESGLOFMDPEEMAV  
FLONAKTLKIKISEASKAFOMEKIEHGYENNHNFTVINRBEKIIREIDFYREDEDE  
EEGGGEGKEEVEGEAVEVEVEVOTEPEDNEPEKASELSOVELOAAAGALP  
VSSPEPPALPPADAPVTOIGFEAPPLQGAAPASGADSEPARHIFSTWLNLSLNE"  
RKATTPPCPTPGSEGLQIQIPPSSEDSNVKAEVAAAASRAVSGKETSAPATISO  
IGFEAPPLQGAAPASGADSEPARHIFSTWLNLSLNE"

FEATURES

source  
CDS  
1. 2098  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_type="cardiomyocyte"  
/tissue\_type="cardiac muscle"  
227. 1873  
/function="putative role in signal transduction"  
/note="60kDa isoform"  
/codon\_start=1  
/evidence="experimental"  
/product="titin zinc-finger anchoring protein"  
/protein\_id="CAC43020.1"  
/db\_xref="GI:14588848"  
/translation="MSASLVNYSFSPKRODMNLEKOLICPLEMFTKPVIIIPCOH  
NIDITKQESTREPKKSDQPCHEHEERINITOLNCEVPCISCKYFGAKHDCQVAP  
LTHVFORKSELSDDGIALIVGSNDVQGVISQEDTCTIECCRKQOECEKEDVL  
YGLIEKRNEMTOVITRTOEELKLEHVALIRKYSDDLLENVSKLVESGLOFMDPEEMAV  
FLONAKTLKIKISEASKAFOMEKIEHGYENNHNFTVINRBEKIIREIDFYREDEDE  
EEGGGEGKEEVEGEAVEVEVEVOTEPEDNEPEKASELSOVELOAAAGALP  
VSSPEPPALPPADAPVTOIGFEAPPLQGAAPASGADSEPARHIFSTWLNLSLNE"  
RKATTPPCPTPGSEGLQIQIPPSSEDSNVKAEVAAAASRAVSGKETSAPATISO  
IGFEAPPLQGAAPASGADSEPARHIFSTWLNLSLNE"

ORIGIN

632 a 497 c 534 g 435 t

Query Match 1.9%: Score 27; DB 9; Length 2098;  
Best Local Similarity 100.0%; Pred. No. 0.0052;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 508 ATCATTCACATCTACAGCAGAGTCC 534  
|||||  
536 ATCATTCACATCTACAGCAGAGTCC 562

RESULT 18  
HSA291712

2202 bp mRNA linear PRI 06-JUL-2001  
HSA291712  
Homo sapiens mRNA for RNF29 gene for ring finger protein 29.

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

AJ291712.1 GI:13160384  
ring finger protein 29; RNF29 gene.  
human.  
Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

SOURCE

human.  
Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

ORGANISM

human.  
Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 2202)  
Centner, T., Yano, D., Kimura, E., McElhinny, A.S., Pelin, K.,  
Witt, C.C., Bang, M.L., Trombitas, K., Granzier, H., Gregorio, C.C.,

TITLE Sorimachi, H. and Labelt, S.  
Identification of muscle specific ring finger proteins as potential  
regulators of the rttin kinase domain  
J Mol. Biol. 306 (4), 717-726 (2001)

JOURNAL MEDLINE 21140140  
PUBMED 11243782  
REFERENCE 2 (bases 1 to 2202)  
AUTHORS Centner, T.  
TITLE Direct Submission  
JOURNAL Submitted (12-FEB-2001) Centner T., Structure and Biocomputing,  
EMBL-Heidelberg, Meyerhofstrasse 1, Heidelberg 69112, Germany  
REMARK Revised by author 22-FEB-2001  
FEATURES Location/Qualifiers  
source 1..2202  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
misc\_feature 133..135  
/note="putative alternative translation initiation codon"  
/evidence="not-experimental"  
gene 181..1779  
/gene="RNF29"  
/note="RNF29"  
/note="alternative"  
/codon\_start=1  
/product="ring finger protein 29"  
/protein\_id="CAC32840.1"  
/db\_xref="GI:13160386"  
/translation="MDNLEKOLICPLEMFKPVYILPCQHNLCRCASDIPQASNP  
YLPTRGGTTMAGSGRCPSCHREYVLDHGYGLGRNLIVENTIDYKQESTRPEK  
SDQPCGHEEERINITYCLNCEVPTSCICVKGALDCQVAPLTHVFRQKSELSDGI  
AIIVGSNDRVGVISQLEDCTIECCCKQKQCELEKEDYLYGLIEERNEMQVIT  
RTQEEKLEHVALIKKYSDDLNLNREKLIIRIDFYREDEDEEEEGEGEKEGSEV  
KAFQMEKLEHGYENNNHFTVNLNREKLIIRIDFYREDEDEEEEGEGEKEGSEV  
GEAVEVEVNVQTFEPGDENDPEKASLSQVLAAPGALPVSSPEPPALPPADA  
PVTOIGFAPPLOGAAPASGASGADSEPARHIFSWNSLNE"  
join(181..1365,1654..1779)  
/gene="RNF29"  
/note="short isoform"  
/codon\_start=1  
/product="ring finger protein 29"  
/protein\_id="CAC32839.1"  
/db\_xref="GI:13160385"  
/translation="MDNLEKOLICPLEMFKPVYILPCQHNLCRCASDIPQASNP  
YLPTRGGTTMAGSGRCPSCHREYVLDHGYGLGRNLIVENTIDYKQESTRPEK  
SDQPCGHEEERINITYCLNCEVPTSCICVKGALDCQVAPLTHVFRQKSELSDGI  
AIIVGSNDRVGVISQLEDCTIECCCKQKQCELEKEDYLYGLIEERNEMQVIT  
RTQEEKLEHVALIKKYSDDLNLNREKLIIRIDFYREDEDEEEEGEGEKEGSEV  
KAFQMEKLEHGYENNNHFTVNLNREKLIIRIDFYREDEDEEEEGEGEKEGSEV  
GEAVEVEVNVQTFEPGDENDPEKASLSQVLAAPGALPVSSPEPPALPPADA  
PVTOIGFAPPLOGAAPASGASGADSEPARHIFSWNSLNE"

BASE COUNT 660 a 489 c 560 g 493 t  
ORIGIN

Query Match 1.9%; Score 27; DB 9; Length 2202;  
Best Local Similarity 100.0%; Pred. No. 0.0052;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 508 ATCATGACATCTACAGCAGAGTCC 534  
|||||  
Db 442 ATCATGACATCTACAGCAGAGTCC 468

RESULT 19 2634 bp mRNA linear PRI 15-JUN-2002  
AC091310  
LOCUS  
DEFINITION Homo sapiens cDNA FLJ33991 fls, clone DFNES2007332, moderately  
similar to Mus musculus RING-finger protein MURF mRNA.  
AC091310  
VERSION AK091310.1 GI:21749650  
KEYWORDS oligo capping; fls (full insert sequence).

SOURCE Homo sapiens normal dermal fibroblasts (Neonatal Skin) (NHDE2564)  
CDNA to mRNA, clone\_lib:DFNES2 clone:DFNES2007332.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S.,  
Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R.,  
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,  
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,  
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,  
Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,  
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,  
Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.  
NEDO human cDNA sequencing project  
TITLE Unpublished  
JOURNAL  
REFERENCE 2 (bases 1 to 2634)  
AUTHORS Isogai, T. and Yamamoto, J.  
TITLE Direct Submission  
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7  
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan  
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; (CDB) full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
key Technology Center etc.); 5'- & 3'- and one pass sequencing: RAB,  
HRI and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: HRI and  
RAB; annotation: HRI and RAB.  
Location/Qualifiers  
source 1..2634  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DFNES2007332"  
/cell\_type="normal dermal fibroblasts (Neonatal Skin)"  
/NHDE2564"  
/clone\_lib="DFNES2"  
/note="cloning vector: PME18SFL3-Primary culture, normal  
dermal fibroblasts"

BASE COUNT 793 a 580 c 586 g 675 t  
ORIGIN

Query Match 1.9%; Score 27; DB 9; Length 2634;  
Best Local Similarity 100.0%; Pred. No. 0.0052;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 508 ATCATGACATCTACAGCAGAGTCC 534  
|||||  
Db 686 ATCATGACATCTACAGCAGAGTCC 712

RESULT 20 121721 bp DNA linear PRI 29-JUN-2002  
AC090195  
LOCUS Homo sapiens chromosome 8, clone RP11-366K18, complete sequence.  
DEFINITION AC090195  
AC090195.5 GI:21629404  
VERSION  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 121721)  
AUTHORS Birren, B., Linton, L., Nussbaum, C., Lander, E., Allen, N., Anderson, S.,  
Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A.,  
Camarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collimore, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J. S.,  
Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,

Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., Lacroque, K., Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Marcquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Mcpheeters, R., Meldrim, J., Menus, L., Mihova, T., Menga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Sougne, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 121721)

Anderson, S., Baran, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhaltier, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S., Ferreira, P., Fitzgerald, M., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., McEwan, P., McKernan, K., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 121721)

Baran, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhaltier, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., McEwan, P., McKernan, K., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT  
Submitted (29-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 29, 2002 this sequence version replaced g1:21592056.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research

## FEATURES

## source

Center code: WIBR  
Web site: <http://www.seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: LI2201  
Center clone name: 366\_K\_18  
-----  
The first 121.7 kb of this clone are being submitted.  
The remainder overlaps accession number AC090196 [WIGR project LI2202].

## Location/Qualifiers

1..121721  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-366K18"  
/clone\_lib="RPC1-11 Human Male BAC"  
470..519  
/rpt\_family="(TG)n"  
624..842  
/rpt\_family="L2"  
complement(1206..1457)  
/rpt\_family="MER34"  
complement(1470..2979)  
/rpt\_family="T199er2"  
complement(2980..3164)  
/rpt\_family="MER82"  
3264..3380  
/rpt\_family="MIR"  
3707..3842  
/rpt\_family="MIR3"  
complement(3937..4400)  
/rpt\_family="MLTIB"  
4253..4258  
/note="<30 qual SNGL region"  
4310..4314  
/note="<30 qual SNGL region"  
4338..4384  
/note="<30 qual SNGL region"  
4950..5208  
/rpt\_family="LTR16A"  
complement(5263..5702)  
/rpt\_family="HAL1"  
complement(5703..5736)  
/rpt\_family="Alu0"  
5737..5776  
/rpt\_family="(TG)n"  
complement(5777..6061)  
/rpt\_family="Alu0"  
complement(6062..6428)  
/rpt\_family="HAL1"  
7448..7707  
/rpt\_family="Alu0"  
complement(8751..8825)  
/rpt\_family="tRNA-Arg-CCG"  
9180..9272  
/rpt\_family="tRNA-Tyr-Trp"  
9801..9889  
/rpt\_family="tRNA-Tyr-Trp"  
10004..10076  
/rpt\_family="tRNA-Ala-GCA"  
11489..11722  
/rpt\_family="MIR"  
11944..12107  
/rpt\_family="L1ME4A"  
12145..12450  
/rpt\_family="Alub"  
12492..13165  
/rpt\_family="L1M4"  
13167..13314  
/rpt\_family="Charlie5"



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 3, 2002, 13:10:09 ; Search time 60 seconds  
(without alignments)  
2349.204 Million cell updates/sec

Title: US-09-908-988B-2  
Perfect score: 1312  
Sequence: 1 MNFVGRKPLIGAHNMNDL.....EGNAGLEERLDVPEGSGIH 366

Scoring table:  
BIOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 6.0, Ygapext 7.0  
Delop 6.0, Delext 7.0

Searched: 341543 segs, 192557720 residues

683086

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n model -DEV=xlh  
Q=/cgn2\_1/USPFO\_spool/US09908988/runat\_26112002.112156.11295/app.query.fasta\_1.519  
-DB=Published.Applications\_NA -OFMT=fastap -SUFFIX=rmpb -MINMATCH=0.1  
-LOOPCL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blissum62  
-TRANS=Numan40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=20 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US0908988.ecgn.1.1.21.gunat.26112002.112156.11295  
-NCPU=6 -ICPU=3 -NO\_XIPYX -NO\_MAP -LARGEDUTRY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1912	100.0	1431	10	US-09-908-988B-1
2	1162	60.8	2590	10	US-09-908-988B-3
3	1147.5	60.0	1762	10	US-09-764-864-34
4	1110	58.1	1597	10	US-09-908-988B-5

*His Appl No 162 not published*  
*At No 913*  
*of 162*

5	1091.5	57.1	1781	10	US-09-764-864-21	Sequence 21, Appl
6	614	32.1	587	10	US-09-764-864-493	Sequence 493, App
7	416	21.8	446	10	US-09-764-864-482	Sequence 482, App
8	280	14.6	1039	10	US-09-880-192-25	Sequence 25, Appl
9	274	14.3	1755	10	US-09-925-300-729	Sequence 729, App
10	234	12.2	2981	10	US-09-925-301-101	Sequence 101, App
11	225	11.8	599	10	US-09-764-864-1594	Sequence 1594, Ap
12	225	11.8	599	10	US-09-764-864-1595	Sequence 1595, Ap
13	225	11.8	599	10	US-09-764-877-2549	Sequence 2549, Ap
14	225	11.8	599	10	US-09-764-877-2550	Sequence 2550, Ap
15	225	11.8	625	10	US-09-764-864-46	Sequence 46, Appl
16	225	11.8	625	10	US-09-764-877-269	Sequence 269, App
17	223.5	11.7	7566	10	US-09-728-952-1	Sequence 1, Appl1
18	200.5	10.5	1332	10	US-09-250-883-14	Sequence 14, Appl
19	199	10.4	3441	10	US-09-880-192-41	Sequence 41, Appl
20	196	10.3	3826	10	US-09-927-091-3	Sequence 3, Appl1
21	191.5	10.0	479	10	US-09-920-300A-1569	Sequence 1569, Ap
22	191.5	10.0	479	12	US-10-033-528-1569	Sequence 1569, Ap
23	186	9.7	3033	12	US-10-044-090-508	Sequence 508, App
24	174.5	9.1	1400	10	US-09-764-864-56	Sequence 56, Appl
25	161	8.4	774	10	US-09-764-864-40	Sequence 40, Appl
26	156	8.2	459	10	US-09-864-761-31683	Sequence 31683, A
27	152	7.9	737	10	US-09-910-943-471	Sequence 471, App
28	151.5	7.9	3035	10	US-09-864-864-311	Sequence 311, Appl
29	143.5	7.5	560	10	US-09-864-761-18057	Sequence 18057, A
30	143.5	7.5	560	10	US-09-864-761-26672	Sequence 26672, A
31	139	7.3	2089	10	US-09-764-864-1598	Sequence 1598, Ap
32	136	7.1	32190	10	US-09-764-878-201	Sequence 201, App
33	135	7.1	45845	10	US-09-927-091-6	Sequence 6, Appl1
34	134.5	7.0	456	10	US-09-864-761-1296	Sequence 1296, Ap
35	134.5	7.0	456	10	US-09-864-761-10030	Sequence 10030, A
36	134	7.0	488	10	US-09-864-761-15156	Sequence 15156, A
37	133	7.0	32249	10	US-09-764-878-202	Sequence 202, App
38	132	6.9	3587	10	US-09-962-436-294	Sequence 294, App
39	132	6.9	3587	10	US-09-880-107-2401	Sequence 2401, App
40	131	6.9	49744	10	US-09-927-091-4	Sequence 4, Appl1
41	130.5	6.8	262	10	US-09-250-883-1	Sequence 1, Appl1
42	130.5	6.8	1310	10	US-09-764-864-1651	Sequence 1651, Ap
43	130	6.8	6421	10	US-09-764-864-1650	Sequence 1650, Ap
44	127	6.6	429	10	US-09-864-761-18056	Sequence 18056, A
45	127	6.6	429	10	US-09-864-761-18059	Sequence 18059, A

#### ALIGNMENTS

RESULT 1  
US-09-908-988B-1  
; Sequence 1, Application US/09908988B  
; Patent No. US20020127690A1  
; GENERAL INFORMATION:  
; APPLICANT: OLSON, ERIC  
; APPLICANT: SPENCER, JEFFREY A.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES  
; FILE REFERENCE: MYOG:0280US  
; CURRENT APPLICATION NUMBER: US/09/908, 988B  
; PRIOR FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: 60/219, 020  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1431  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (199)..(1296)  
US-09-908-988B-1  
Alignment Scores: 1.51e-210 Length: 1431  
Pred. No.: 1912.00 Matches: 366  
Score:

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0  
US-09-908-988b-2 (1-366) x US-09-908-988b-1 (1-1431)  
QY 1 MetaspheThValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
DB 199 ATGAACCTTCACGGTGGTTTCAAGCCGCTCTAGGGGATGGCCACACATGACACTTG 258  
QY 21 GluLysGlnLeuLeuIleCysProIleCysLeuGluMetPheSerLysProValIleLeu 40  
DB 259 GAGAGCAGCTCATTTGGCCCCATCTGCTGAGATGTTCTCAAGCCCGTGATCTTG 318  
QY 41 ProCysGlnHisAsnLeuLysAsnArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
DB 319 CCGTCCACACACACCTGTGGCCGAGTGTGCCACAGCAGCTCTTCCAGGCCCTCAATCT 378  
QY 61 LeuTrpGlnSerArgGlySerThrThrValSerSerGlyArgPheArgCysProSer 80  
DB 379 CTGTGGCAATCCCGGGGCTCCACACGGTGTCTTCAGAGAGAGCTTCCGATGCCATCT 438  
QY 81 CysArgHisGluValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100  
DB 439 TGTAGGCACAGAGTGTCTCTGGACAGCATGTGTCTATGAGCTGCAGCGGACCTGCTA 498  
QY 101 ValGluAsnIleIleAspIleTyrLysGlnGluSerSerArgProLeuHisAlaLysAla 120  
DB 499 GTGGAGACATATTGTACATCTTACAAAGCAGAGAGTCTCCGGCCACATGCACGCCAGGCT 558  
QY 121 GluGlnHisLeuMetCysGluGlnHisGluAspGluLysIleAsnIleTyrCysLeuSer 140  
DB 559 GACACGACACTCATGTGTGTAGAGAGCAGAGAGAGAGACATCAATCTACTGCTGAGC 618  
QY 141 CysGluValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGluVal 160  
DB 619 TGGCGAGTGGCCACCTGCTCTCTGCAAGGTTTGGCGCCCCACAGAGACTGTGAGGTG 678  
QY 161 AlaProLeuProThrIleTyrLysArgGlnLysSerGluLeuSerAspGlyIleAlaMet 180  
DB 679 GCCCTCTGCCCCACATTTTACAAACGCCAGAAAGATGAGCTGACGATGGCATGGCAGATG 738  
QY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGluGluValCysGln 200  
DB 739 CTGGTGGGGGCAATGACCGTGTGTGCGAGCATATCACCCAGATGGAGAGGTGTGCCAG 798  
QY 201 ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGluThrLeu 220  
DB 799 ACCATTGAGGACACACAGCCGACAGAGCAAGCACTGTTAAACGAGAGTTCCGACACCTG 858  
QY 221 CysAlaValLeuGluGlnArgLysGlyGluLeuLeuGlnAlaLeuAlaArgGluGln 240  
DB 859 TGGCGGATTTTGGAGAGGCGCAAGGCGAACTCTTCAACACACTGGCCCGGACAGAGAG 918  
QY 241 GluLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyLysPheHisLeuGlySer 260  
DB 919 GAGAGATTGACAGCGGTGGGGCTCATTCGCCAGAGGAGACCACTTGGAGGGCTCC 978  
QY 261 SerLysLeuValGluSerAlaIleGlnSerMetGluGluProGlnMetAlaLeuTyrLeu 280  
DB 979 TCAAAAGCTGGTAGAGTCCGCCATCCAGTCCATGGAGAGCGCGAGATGGCTCTACTC 1038  
QY 281 GlnGlnAlaLysGluLeuIleAsnLysValGlyAlaMetSerLysValGluLeuAlaGly 300  
DB 1039 CAGCGGCAAAAGAGCTGTATCAACAAGGTCGGGGCAATGTCGAAGGTGGAGCTGGCAGGA 1098  
QY 301 ArgProGluProGlyTyrGluSerMetGluGlnPheSerValSerValGlnHisValAla 320  
DB 1099 CGGCCGGAGCCAGGCTATGAGCATGAGCAATTTCTCTGTGAGCGGTGGAGCAGCTGGCC 1158  
QY 321 GluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGluGluAspAsp 340

DB 1159 GAAATGTTGGCAACATGACTTCCAGCCGGCCGCTGGGGATGAAGAGATGACAGC 1218  
QY 341 MetAlaLeuAspGlyGluGluGlyAsnAlaGlyLeuGluGlnArgLeuAspValPro 360  
DB 1219 ATGGCTTTGGATGGGAGAGGCGCAATGCGGGCTGGAGAGAGCGGCTGGACGTGCCA 1278  
QY 361 GluGlySerGlyLeuHis 366  
DB 1279 GAGGCTCAGCGCTGCAC 1296  
RESULT 2  
US-09-908-988b-3  
; Sequence 3, Application US/0908988B  
; Patent No. US20020127690A1  
; GENERAL INFORMATION:  
; APPLICANT: OLSON, ERIC  
; APPLICANT: SPENCER, JEFFREY A.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES  
; FILE REFERENCE: MYOG:028US  
; CURRENT APPLICATION NUMBER: US/09/908, 988B  
; PRIOR FILING DATE: 2000-07-18  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2590  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (80)..(1714)  
US-09-908-988b-3  
Alignment Scores:  
Pred. No.: 7,37e-124 Length: 2590  
Score: 1162.00 Matches: 226  
Percent Similarity: 76.99% Conservative: 55  
Best Local Similarity: 61.92% Mismatches: 76  
Query Match: 60.77% Indels: 8  
Gaps: 4  
US-09-908-988b-2 (1-366) x US-09-908-988b-3 (1-2590)  
QY 1 MetaspheThValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
DB 80 ATGAGCACTTCTGTGAATTTCAAACTTTTCTCAAAAGCAGACAGACCATGGATACTTG 139  
QY 21 GluLysGlnLeuLeuIleCysProIleCysLeuGluMetPheSerLysProValIleLeu 40  
DB 140 GAAAGCACTGTATCTTCCATCTGCTAGAGATGTTCCAGAAACCTGTGGTCAATTC 199  
QY 41 ProCysGlnHisAsnLeuLysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
DB 200 CTTTGGCAGACAAACCTGTGCGAGAAATGTGCCAGTGCATCTTCCAGGCTCTTAACCG 259  
QY 61 LeuTrpGlnSerArgGlySerThrThrValSerSerGlyArgPheArgCysProSer 80  
DB 260 TACTTACCCACAGAGAGAGGACACACCGTGGCATTCAGAGGGGCGGCTGCTCCCTCC 319  
QY 81 CysArgHisGluValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100  
DB 320 TGCAGACATAGAGTGTGTGAGACAGACATGGGGCTGTATGGACTGCAAGAACCTGTCTC 379  
QY 101 ValGluAsnIleIleAspIleTyrLysGlnGluSerSerArgProLeuHisAlaLysAla 120  
DB 380 GTGGAAACCTTTATGTATCTACAGCAGAAATCCACAGGCCA---GAAAAAATTG 436  
QY 121 GluGlnHisLeuMetCysGluGlnHisGluAspGluLysIleAsnIleTyrCysLeuSer 140  
DB 437 GACACAGGCC---ATGTGTGAAGAGCATGAAAGAGAAACCATCAACATGTATGTGTGAAC 493

QY 141 CysgluValProThrcysSerLeucCysLysValPheGlyAlaHisLysAspCysGluVal 160  
 |||||  
 DB 494 TGTGAAGTCCACCGTCTCTTGTGCAAGGTTTTTGCCGCCCATAGAGACTGCCAGTGC 553  
 |||||  
 QY 161 AlaProLeuProThrlleTyrllysArgGlnLysSerGluLeuSerAspGlyIleAlaMet 180  
 |||||  
 DB 554 GCTCCCTGACTACATGTGTTCAGAGAGAGAAAGTCAAGAGTCAAGTGTATGTTGCTGTA 613  
 |||||  
 QY 181 LeuValAlaIleAspArgValGlnAlaValIleThrcLmetGluGluValCysGln 200  
 |||||  
 DB 614 CTGTGGGAGAGACAGATAGATGCCAGGTGTGATCAGCCAGCTGGAGGACCTGTAA 673  
 |||||  
 QY 201 ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGluThrLeu 220  
 |||||  
 DB 674 ACTATGTAGAGAGTGTGTCAGAAAGAGAAACAGAGACTGTGTGAGAAATTTGATCAGCTA 733  
 |||||  
 QY 221 CysAlaValLeuGluGluArgLysGlyGluLeuLeuGlnAlaLeuAlaArgGluGlu 240  
 |||||  
 DB 734 TACGGCATCTCTGGAGAGAGAGAGAGACTGAATGACCCAGCTCACTCAGACACAGAG 793  
 |||||  
 QY 241 GluLysLeuGlnArgValArgGlyLeuIleArgLysIleArgLysPheLysGluLysSer 260  
 |||||  
 DB 794 GAGAACTGACACATGTCCGAGACTCTTATCAGAGAACTATTCGATCCAGCTGGAGACGTA 853  
 |||||  
 QY 261 SerLysLeuValGluSerAlaIleGlnSerMetGluGluProGlnMetAlaLeuTyrlleu 280  
 |||||  
 DB 854 TCCAGATGTGTGATGATCAGAAATCCAGTTCATGATGAGCCCAAAATGGCAGATTTCTG 913  
 |||||  
 QY 281 GlnGlnAlaLysGluLeuLysLysValGlyAlaMetSerLysValGluLeuAlaGly 300  
 |||||  
 DB 914 CAGAAATGCCAACAACCTGTGCAAAAGATCGTGGAAAGATCAAAAGCGTTTCAGATGAG 973  
 |||||  
 QY 301 ArgProGluProGlyTyrlleuSerMetGluGlnPheSerValSerValGlnHisValAla 320  
 |||||  
 DB 974 AAACATAGAACAGATTATGATGATCATGAGCACTTCACTGTCAATCTCATAGAGAA 1033  
 |||||  
 QY 321 GluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGluLysAspAsp 340  
 |||||  
 DB 1034 AAAATATCTCCGAAATGACCTT-----TCTAGAGAAAGAGAGAGAAAGAT 1084  
 |||||  
 QY 341 MetAlaLeuAspGlyGluGluGlnLysAlaGlyLeuGluGlnLysArgLeuAspValPro 360  
 |||||  
 DB 1085 GCAGAGAAATAGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1135  
 |||||  
 QY 361 GluLysSerGlyLeu 365  
 |||||  
 DB 1136 GAGCGAGAAATGTT 1150  
 |||||  
 RESULT 3  
 US-09-764-864-34  
 : Sequence 34, Application US/09764864  
 : Patent No. US20020132753A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Rosen et al.  
 : TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 : FILE REFERENCE: P1223  
 : CURRENT APPLICATION NUMBER: US/09/764,864  
 : PRIORITY FILING DATE: 2001-01-17  
 : Prior application data removed - consult PALM or file wrapper  
 : NUMBER OF SEQ ID NOS: 1792  
 : SOFTWARE: PatentIn Ver. 2.0  
 : SEQ ID NO 34  
 : LENGTH: 1762  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : US-09-764-864-34  
 Alignment Scores:  
 Pred. No.: 1,92e-122 Length: 1762  
 Score: 1147.50 Matches: 225  
 Percent Similarity: 75.34% Conservative: 50  
 Best Local Similarity: 61.64% Mismatches: 72  
 Query Match: 60.02% Indels: 18

DB: 10 Gaps: 4  
 US-09-908-988b-2 (1-366) x US-09-764-864-34 (1-1762)  
 QY 1 MetAspPheThValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
 |||||  
 DB 157 ATGACGCGATCTCTGAATATTACAATCTTTTCCAAAGACAGACAGACCATGGATTAACCTTA 216  
 |||||  
 QY 21 GluLysGlnLeuIleCysProIleCysLeuGluMetPheSerLysProValAla-Ilele 40  
 |||||  
 DB 217 GAGAACTACATCATCTGCCATCTGTAGAGATGTTCCAGAAACCTGTGTGATGATTTCT 276  
 |||||  
 QY 40 uProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPr 60  
 |||||  
 DB 277 CCTTGTGACCAACCTGTGTGAGAAATGTCCAGTGAATTTTCCAGGCTCTTAACCC 336  
 |||||  
 QY 60 OLeuThrPheLysSerArgLysSerThrValSerSerGlyArgPheArgCysProse 80  
 |||||  
 DB 337 GTATTGTCCCAACAAGAGAGATGATCCACCATGCGATCAGGGGCGGATTCGGTGGCCCATC 396  
 |||||  
 QY 80 rCysArgHisGluValValLeuAspArgHisGlyValTyrlleGluGlnArgAsnLeu 100  
 |||||  
 DB 397 CTGTAGACATGAAAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 456  
 |||||  
 QY 100 uValGluAsnIleIleAspIleTyrlleGlnLysGlnLysSerArgProLeuHisAlaLysAl 120  
 |||||  
 DB 457 GGTGAAATATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 510  
 |||||  
 QY 120 aGluGlnHisLeuMetCysGluGlnHisGluLysAspGlyLysIleAsnIleTyrlleCysLeu 140  
 |||||  
 DB 511 ATCCGACCAAGCCATGTCGAGAGACATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570  
 |||||  
 QY 140 rCysGluValProThrcysSerLeuCysLysValPheGlyAlaHisLysAspCysGluVal 160  
 |||||  
 DB 571 CTGGAAATGATCCACCTCTCTCTGTGCAAGGTGTGTGTGCAACAAGAGAGAGAGAGAGAG 630  
 |||||  
 QY 160 AlaProLeuProThrlleTyrllysArgGlnLysSerGluLeuSerAspGlyIleAlaMet 180  
 |||||  
 DB 631 GGTGTCCTCTCACTCATGTGTTCACAGAGACAGAAAGTGTGAGTCAAGTATGATGATGATG 690  
 |||||  
 QY 180 tLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGluValCysGly 200  
 |||||  
 DB 691 CCTGTGGCGACAGCAAGATCGAGTCCAGAGAGATGATCCAGCTGGAGAGACACTGCA 750  
 |||||  
 QY 200 nThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGluThrle 220  
 |||||  
 DB 751 AACATGAGAGAAATGTTCCAGAAACAGAAACAGAGAGCTTGTGAGAAAGTGTGATTAACCT 810  
 |||||  
 QY 220 uCysAlaValLeuGluGlnLysArgLysGlyGluLeuLeuGlnAlaLeuAlaArgGluGln 240  
 |||||  
 DB 811 GTATGGCATTTTGGAG 870  
 |||||  
 QY 240 uGluLysLeuGlnArgValArgGlyLeuIleArgGlnTyrlleGlyAspHisLeuGluLys 260  
 |||||  
 DB 871 GGAGAACTGGAACATGTCCTGCTGTATCAAAAAGATTTCTGATCATTTGTGAGAAACCT 930  
 |||||  
 QY 260 rSerLysLeuValGlnSerAlaIleGlnSerMetGluGlnProGlnMetAlaLeuTyrlle 280  
 |||||  
 DB 931 CTCAAAAGTGTGTGAGTCAAGAAATCAAGTGTGTGATGATGATGATGATGATGATGATGAT 990  
 |||||  
 QY 280 uGlnGlnAlaLysGluLeuLysLysValGlyAlaMetSerLysValGluLeuAlaGly 300  
 |||||  
 DB 991 GCAGAAATGCCAAACCCCTGTAATAAATAATCTCAAGACATCAAAAGGCGATTTCCAGATGGA 1050  
 |||||  
 QY 300 yArgProGluProGlyTyrlleGlnSerMetGluGlnPheSerValSerValGlnHisValAl 320  
 |||||  
 DB 1051 GAAATAGAAATGCTGATGAGACATGAAACCACTTCAAGTCAACCTCAATATAGAGAGAGA 1110  
 |||||  
 QY 320 aGluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGluLysAspAsp 340  
 |||||  
 DB 1111 AAAGATATATGCTGAATTTGACTTTTAC-----AGAGAGAGAGAGAGAGAGAGAGAGAG 1152  
 |||||  
 QY 340 pMetAlaLeuAspGlyGluGlnLysAlaGlyLeuGluGlnLysArgLeuAspValPr 360

Db 1153 T-----GAAGAAGAGAGAGCGGAGAGAGAGAAAG----- 1186  
QY 360 oGluGlyserGly 364  
Db 1187 -GAAGAGAGAGCA 1198  
RESULT 4  
US-09-908-988b-5  
; Sequence 5, Application US/09908988B  
; Patent No. US20020127690A1  
; GENERAL INFORMATION:  
; APPLICANT: OLSON, ERIC  
; APPLICANT: SPENCER, JEFFREY A.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES  
; FILE REFERENCE: MYOG:02805  
; CURRENT APPLICATION NUMBER: US/09/908,988B  
; CURRENT FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: 60/219,020  
; PRIOR FILING DATE: 2000-07-18  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 1597  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (299)..(1327)  
US-09-908-988b-5  
Alignment Scores:  
Pred. No.: 3,41e-118 Length: 1597  
Score: 1110.00 Matches: 212  
Percent Similarity: 78.53% Conservative: 55  
Best Local Similarity: 62.35% Mismatches: 63  
Query Match: 58.05% Indels: 10  
DB: 10 Gaps: 3  
US-09-908-988b-2 (1-366) x US-09-908-988b-5 (1-1597)  
QY 4 ThrValGlyPheLys-----ProLeuGluGlyAspAlaHisAsn 16  
Db 278 ACAGTCGGATTTCAAGCAATATGATTAATATAGCTGATCTCTGATGAAAGCGCT 337  
QY 17 MetAspAsnLeuGluGlyGlnLeuIleCysProIleCysLeuGluMetPheSerLysPro 36  
Db 338 ATGGAGAACCTGGAGAGACGATGATCTGCCCATCTGCTGGAGATGTTTACCAAGCGCT 397  
QY 37 ValValIleLeuProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGln 56  
Db 398 GTGGTATCTCTGCTGCTGCCACACACCTCTGCCGAGAGTGTGCCAAGCAGATCTTCCAG 457  
QY 57 AlaSerAsnProLeuTrpGlnSerArgGlySerThrValSerSerGlyArgPhe 76  
Db 458 GCTGGAGATTCCTACTGACCAACCGCGGTGCTCACTGCTGCGAGTCTTC 517  
QY 77 ArgCysProSerCysArgHisGluValValLeuAspArgHisGlyValIleGlyLeuGln 96  
Db 518 CGTGGCCCTCTGCTGCCCATGAGATCATGACCGGCGGAGGTGTAGCGCTGAG 577  
QY 97 ArgAsnLeuLeuValGluAsnIleIleAspIleTyrLysGlnGlu---SerSerArgPro 115  
Db 578 AGGAACCTGCTGCTGAGAAACATCATTCATTCACAGAGAGAGTCTCCAGTCGGGCC 637  
QY 116 LeuHisAlaLysAlaGluGlnHisLeuMetCysGluGlnHisGluAspGluLysIleAsn 135  
Db 638 CTG-----CAAAAGGAGGAGCCAGCGGATGTGCAAGAGAACAGAGAGATCAAC 691  
QY 136 IleTyrCysLeuSerCysGluValProThrCysSerLeuCysLysValPheGlyAlaHis 155  
Db 692 ATCTACTGTCTACGCTGTGAGGTGCTTACTTGTCTCTTGTGCAAGGTGTGTGGGCTCAC 751

QY 156 LysAspCysGluValAlaProLeuProThrIleTyrLysArgGlnLysSerGluLeuSer 175  
Db 752 CAGGCTGTGAGGTGCTCTTTCAGAAAGCATCTTCCAGAGCAGAAAGACTGAGCTGAGT 811  
QY 176 AspGlyIleAlaMetLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMet 195  
Db 812 AACGTGATCTCCATGCTGTGGTGGGGGAGACGACGAGTGCAGAGATCATCTCTCAGCTG 871  
QY 196 GluGluValCysGlnThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGln 215  
Db 872 GAGACTGCTGACAGTGCACCAAGAGAAATAGCCACCGAGGAGAGAGAGACTGCTGAG 931  
QY 216 ArgPheGluThrLeuCysAlaValLeuGluGluArgLysGlyGluLeuLeuAlaLeu 235  
Db 932 AGTTTGACACCTCTTACGCGCATCTTGATGAGAGAGAGACGAGCTCTGACGGGATC 991  
QY 236 AlaArgGluGlnGluGluLysLeuGlnArgValArgGlyLeuIleArgGlnIleGlyLysP 255  
Db 992 AGCAGAGAGAGAGAGAGAGAGAGTGGGCTTCAGAGCTGTGATCTTCACAGAGAG 1051  
QY 256 HisLeuGluGlySerSerLysLeuValGluSerAlaIleGlnSerMetGluLysProGln 275  
Db 1052 CAGCTGAAAAGTCCACCAAGCTTGTGAGACCCGCAATCCAGTCCGTGATGAGCCGGA 1111  
QY 276 MetAlaLeuTyrLeuGlnGlnAlaLysGluLeuIleAsnLysValGlyAlaMetSerLys 295  
Db 1112 GGGGCTACTCTCTCTTCACAGTCCACAGCAGCTCATCAAGACATTTGTGAAGCCCTCAAG 1171  
QY 296 ValGluLeuAlaGlyArgProGluProGlyTyrGluSerMetGluGlnPheSerValSer 315  
Db 1172 GCGTGCAGCTGGGAGAGAGAGAGAGCGCTTGTGAGAAATGATGACTTACTCTGAGC 1231  
QY 316 ValGluHisValAlaGluMetLeuArgThrIleAspPheGlnProGlyAlaAlaLysP 335  
Db 1232 TTAGAACAATATACAGAGGCGCTTGAAGGCGCATGACTTGTGGAGACGTTAAAGATGTGAT 1291  
RESULT 5  
US-09-764-864-21  
; Sequence 21, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 21  
; LENGTH: 1781  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-864-21  
Alignment Scores:  
Pred. No.: 5,41e-116 Length: 1781  
Score: 1091.50 Matches: 209  
Percent Similarity: 77.71% Conservative: 63  
Best Local Similarity: 59.71% Mismatches: 67  
Query Match: 57.09% Indels: 11  
DB: 10 Gaps: 4  
US-09-908-988b-2 (1-366) x US-09-764-864-21 (1-1781)  
QY 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluLysGlnLeuIleCysProIleCys 29  
Db 137 CTGATCCAGATGAGGAAATCCATGAGAACTTGGAGAGAGAGCTGATCTGCCCTATCTGC 196  
QY 30 LeuGluMetPheSerLysProValValIleLeuProCysGlnHisAsnLeuCysArgLys 49  
Db 197 CTGAGATGTTTACCAAGCAGGTGATCTTGTCCGTCCAGACACCTGTGCGGAG 256



QY	50	CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGlnSerArgGlySerThrThr	69
Db	257	TGTGCGCAATGACATCTTCTCCAGGCTGCAAAATCTCTACTTGACACCGGGGACGTCACAG	316
QY	70	ValSerSerGlyLysArgPheArgCysProSerCysArgHisGluValValLeuAspArg	89
Db	317	TCCATCTCTGGAGGCGGCTTCCGCTGCCACACCTGCGCCACGAGGTATCATGATGATCGT	376
QY	90	HisGlyValTyrGlyLeuGlnArgAsnLeuLeuValGluAsnIleIleAspIleTyrLys	109
Db	377	CACGGAGTTTAAAGCGCTCGACAGAACCTCTGGTGGAGAACATCATGACATCTTACAA	436
QY	110	GlnGlu---SerSerArgProLeuHisAlaLysAlaGluGlnHisLeuMetCysGluGln	128
Db	437	CAGAGTGTCTCCAGTGGCGCGCTG-----CAGAAAGGCACTCAACCCCATGTGCAAGAG	490
QY	129	HisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThrCysSerLeu	148
Db	491	CACGAGATGAGAAAATTAACATCTACTGTCTCCAGTGTGAGGTGCCACCTGCTCCATG	550
QY	149	CysLysValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIleTyrLys	168
Db	551	TCCAAAGTGTGTGGATCCACAAAGCGCTGGAGGTGGGCCCATTCGACAGTGTCTTCCAG	610
QY	169	ArgGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgVal	188
Db	611	GGACAAAACACTGAACCTGATACTGATCTTCATCTGCTGGTGGCGGGGAATGACCGTGTG	670
QY	189	GlnAlaValIleThrGlnMetGluGluValCysGlnThrIleGluAspAsnSerArgArg	208
Db	671	CAGACCATCATCACTACGCTGAGAGATTCGCCGTGATGACCAAGAGAAAGATCACCAG	730
QY	209	GlnLysGlnLeuLeuAsnGlnArgPheGluThrLeuCysAlaValLeuGluGluArgLys	228
Db	731	GTAAGGGAAGACTGTGAGCCAGAAAGTTTGACACGCTTGTATGCTCATCTGGATGAGAA	790
QY	229	GlyGluLeuLeuGlnAlaLeuAlaArgGluGlnGluLysLeuGlnArgValArgGly	248
Db	791	AGTGAATCTCTCAGCGGGATCAACGCAAGGACAGAGAAAAGCTTACCTTCATGAGGCC	850
QY	249	LeuIleArgGlnTyrGlyAspHisLeuGluGlySerSerLysLeuValGluSerAlaIle	268
Db	851	CTCACTCACACAGTACACAGAGACGCTGGACAACTCCACAAAGCTGGTGGAACTGCCATC	910
QY	269	GlnSerMetGluGluProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGluLeuIleAsn	288
Db	911	CAGTCCCTGGACGAGGCTGGGGGAGCCACTTCTCTGACTGCCAAGCAACTCATCAAA	970
QY	289	LysValGlyAlaMetSerLysIleGluLeuAlaGlyArgProGluProGlyTyrGluSer	308
Db	971	AGCATTTGTGGAAGCTTCCAAAGGCGCTCCACCTGGGGAGACACAGACAGGCGTTTGAGAC	1030
QY	309	MetGluGlnPheSerValSerValGluHisValAlaGluMetLeuArgThrIleAspPhe	328
Db	1031	ATGCACTTCTTACTTGTGATTTAGAGACATATGACAGCCCGTGAAGCATTTGACTTT	1090
QY	329	GlnProGlyAlaAlaGlyAspGlnGluAspAspMetAla-----LeuAsp	344
Db	1091	-----GGACACATATGAGAGAGGAAGAAATTCATTGAAAGAAAGATCAGGAA	1138
QY	345	GlyGluGlnGlyAsnAlaGlyLeuGlnGlu	354
Db	1139	GAGGAAGAGTCCACAGAAAGGAAGAAAGAA	1168

RESULT 6

US-09-764-864-493

; Sequence 493, Application US/09764864

; Patent No. US20020132753A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PT123

QY	218	GLuThrlencuysAlaValleuGluuIuArGlySGlYglIuuDeuGlnlAlaValuIaArG	233
Db	4	GACACGTTGTATGCGCATCTCGGATGAGAAAGAAAGATGAGTGTCTGCAGCGGATCACAGCAG	63
QY	238	GLuGlnGluGluysLeuGlnArGValArGlyLeuIleArGlnIuTyrlGlyAspHisLeu	257
Db	64	GAGAGAGAGAAAGATCTTACTTCATGAGAGGCCCTCATCCAGACAGATACAGAGAGCAGCTG	123
QY	258	GLuGlySerSerIysLeuValIuGSerAlaIleGlnSerMetGluIuProGlnMetAla	277
Db	124	GACAAAGTCCACAAAGCGGTGGGAACATGCATCACTCCCTGGACAGACCTGGGGAGCC	183
QY	278	LeuTyrlencuGlnAlaIysGluLeuIleAsnIysValGlyAlaMetSerIysValGlu	297
Db	184	ACCTCTCTCTTGAAGTGCACAGCAACTCATTAAGATATGTGGAACCTTCAGAGGCTGC	243
QY	298	LeuAlaGlyArGProGluProGlyTyrlGluSerMetGluGlnIuPheSerValSerValGlu	317
Db	244	CAGCTGGGAGAGACAGACGAGCGGCTTTGAAACATGAGACTTCTTACTTGGATTTAAG	303
QY	318	HisValIleAlaGluMetLeuArGThrIleAspPheGlnProGlyIleAlaGlyAspGluGlu	337
Db	304	CACATYACCAAGCGCCCTGGAAGCCATGACTT-----GGACACATGAGGAA	351
QY	338	AspAspAspMetAla-----LeuAspGlyGluGlnGlyAsnAlaGlyLeuGlu	353
Db	352	GAGAGACAAATTCATTGAAAGAAAGATCAGAGAAAGAGAAAGATCCACAGAAAGGAGGAA	411
QY	354	GLu 354	
Db	412	GAA 414	
RESULT 9			
US-09-925-300-729	Sequence 729, Application US/09925300		
US-09-925-300-729	Patent No. US20020151681A1		
US-09-925-300-729	GENERAL INFORMATION:		
US-09-925-300-729	APPLICANT: Craig Rosen,		
US-09-925-300-729	APPLICANT: Steve Ruben		
US-09-925-300-729	TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies		
US-09-925-300-729	FILE REFERENCE: PA101		
US-09-925-300-729	CURRENT APPLICATION NUMBER: US/09/925, 300		
US-09-925-300-729	CURRENT FILING DATE: 2001-08-10		
US-09-925-300-729	PRIOR APPLICATION NUMBER: PCT/US00/05988		
US-09-925-300-729	PRIOR FILING DATE: 2000-03-08		
US-09-925-300-729	PRIOR APPLICATION NUMBER: 60/124, 270		
US-09-925-300-729	PRIOR FILING DATE: 1999-03-12		
US-09-925-300-729	NUMBER OF SEQ ID NOS: 1690		
US-09-925-300-729	SOFTWARE: PatentIn Ver. 2.0		
US-09-925-300-729	SEQ ID NO 729		
US-09-925-300-729	LENGTH: 1755		
US-09-925-300-729	TYPE: DNA		
US-09-925-300-729	ORGANISM: Homo sapiens		
Alignment Scores:			
QY	17	MetAspAsnLeuGluGlnLeuIleGlySerIleGlySerLeuGluMetPheSerIysPro	36
Db	200	ATGGAGCTGCTTGAAGAAGATCTACATGCCCCATTTGTGTAGTCTGTGTGATGATCA	258
QY	37	ValValIleLeuProCysGlnHisAsnLeuGlySerGlyCysAlaAsnAspValPheGln	56
Db	260	CGGTT---TTGGCTGTGCTCCACACACTTTCGCAAAAAATGCTTAGAAGATGTTAGAA	316

```

QY 57 Alaser-----AsnProLeuTrpGlnSerArgGlySerThrValSerSerglyGly 74
Db 317 GGGAGTGTGGGAATTCCTGTGGAGACCA-----GCT 349
QY 75 ArgPheArgCysProSerCysArgHisGluValValLeuAspArgHisGlyValTyrGly 94
Db 350 CCATTCAAGTGTCTCAATGCGTGAAGAAACT-----TCAGCTACTCGTAATTAATAGC 403
QY 95 LeuGlnArgAsnLeuLeuValGluAsnIleLeuAspIleTyrLeuGlnGlnSerSerArg 114
Db 404 CTGCAAGGTAAATTAATCTCCCTGAGGATATGTGGAAAGTATATACAGATCAAGATCTCT 463
QY 115 ProLeuHisAlaValAlaGluGlnHisLeuMetCysGluGlnHisGluAspGluTyr 134
Db 464 CCC-----AAAATGCCAGTATGCAAAAGACACTTGGGGCAGCCCTCTC 505
QY 135 AsnIleTyrCysLeuSerCysGluValProThrCysSerLeuCysLysValPheGlyAla 154
Db 506 AACATTTCTGTGCTGACTGATGATGCAAGTGAATTTGTGGATGTGCTGCTACTCTGGGGAG 565
QY 155 HisLysAspCysGluValAlaProLeuProThrIleTyrLysArgGlnLysSerGluLeu 174
Db 566 CACACCAACATGTCTTCTGTCTTATGAGATGCTATGCTCAGGAAG-----616
QY 175 SerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGln 194
Db 617 -----GATGCTTTGAGTCCCTCCAGAGC 643
QY 195 MetGlu-----GluValCysGlnThrIleGluAspAsnSer 206
Db 644 TTTCAGCTGGCTGCGGGAGAGATGCTTCTGCTGCTGATGATGCAATTCGTAAG 703
QY 207 ArgArgGlnLysGlnLeuLeu-----AsnGlnArgPheGluThr 219
Db 704 AGGAATATCCCTACAGTACTGACTAAAGATTCAAGATTAAGTGAAGAAATTTTGGAGAG 763
QY 220 LeuCysAlaValLeuGlnLysArgLysGlyGluLeuLeuGln-----233
Db 764 TTACAAACACACCTGATCAAAAGAAATGAAATCTGCTGACTTGGAGACCATGAA 823
QY 234 ---AlaLeuAlaArgGlnGlnGluLysLeuGlnArgValArgGlyLeuIleArgGln 252
Db 824 CTGGCTGTATATGACATATGATGACCAAGATCAACAACTCAACACATCTGCAAGAG 883
QY 253 TyrGlyAspHisLeuGlnGlySerSerLysLeuValGlnSerAlaIleGlnSerMetGlu 272
Db 884 -----CAACGGATGGCTTTAACATGCTGAGGCTTTCAAAAGATGTGTCA 928
QY 273 GluProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGluLeuLeuAsnLysValGlyAla 292
Db 929 GAACCC-----ATTGTATTTCTGTGACACAGATGCAAGAGTTTGAAGAAATCAAAAGTA 982
QY 293 MetSerLysValGlnLeuAlaGlyArgProGluProGlyTyrGlnSerMetGluGlnPhe 312
Db 983 ATCAAGGAACCTCTTACCTCCCTTAATTTGCTGCAAGCCCTTAATGAGAAGACTTT 1042
QY 313 SerValSerValGlnHisValAlaGluMetLeuArgThrIleAspPheGlnProGlyAla 332
Db 1043 GATACCACT-----1051
QY 333 AlaGlyAspGlnGluAspAspAspMetAlaLeuAspGlyGluGlnLysAlaGlyLeu 352
Db 1052 -----CACTGGGAACACATTAATACTA-----GTC 1075
QY 353 GluGlnGluArgLeuAspValProGluGlySergly 364
Db 1076 GATGTGATTAACCTTTTTCCTCAAGACACTGCG 1111

```

```

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 101
; LENGTH: 2981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (293)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-101

Alignment Scores:
Pred. No.: 5,48e-17 Length: 2981
Score: 234.00 Matches: 80
Percent Similarity: 41.95% Conservative: 58
Best Local Similarity: 24.32% Mismatches: 139
Query Match: 12.24% Indels: 52
DB: 10 Gaps: 13

US-09-908-988b-2 (1-366) x US-09-925-301-101 (1-2981)
QY 20 LeuGlnLysGlnLeuIleLeuAspProIleCysLeuGlnMetPheSerLysProValIle 39
Db 366 CTGCAGCAGGAGACCACTGCCCCGTGCTGCTGCAAGTACTTGCAGAGCC---ATGATG 422
QY 40 LeuProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnIleSerAsn 59
Db 423 CTCGACTGCGGCACAAATCATGTTGGCGGTGCTC-----GGC 461
QY 60 ProLeuTrpGlnSerArgGlySerThrThrValSerSerglyGlyArgPheArgCysPro 79
Db 462 CGCTGCTGCG---GGCAGCGCAGAGACTAACGCTGCG-----TGCCTCG 500
QY 80 SerCysArgHisGlnValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 99
Db 501 CAGTCCCG---GAGACCTTCCCGCAGAGCCAC-----ATGGGCCCAACCGG 545
QY 100 LeuValGluAsnIleIleAspIleTyrLysGln---GluSerSerArgProLeuHisAla 118
Db 546 CACCTGGCCAACTGATCCCACTGTAAGCAAGCTGCGCAGCGGCGCTCGGGCGCC 605
QY 119 LysAlaGlnGlnHisLeuMetCysGlnGluHisGluAspGluLysIleAsnIleTyrCys 138
Db 606 GCGCGCAGATGGGC---GTGTGCGAGAGACCGCC---GAGCCCTTGAAGCTGTACTGC 659
QY 139 LeuSerCysGluValProThrCysSerLysLysValPheGlyAlaHisLysAspCys 158
Db 660 GAGGAGACCAAGATGCAATCTGCTGTGTGTCGACCGCTCCCGGAGACCGGGCCAC 719
QY 159 GluValAlaProLeuProThrIleTyrLysArgGlnLysSerglyLeuSerAspGlyIle 178
Db 720 AGCGTGTGCTGCGCTGAGAGAGCGGTGAGGGCTTCAAGACCAATTCAGAAACGATC 779
QY 179 AlaMetLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGluGluVal 198
Db 780 GACCATTTAAAGAGTGAAGATTAAAGAGAGACGTCGGCGCAGGGGAGACAGCA 839
QY 199 CysGlnThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlu 218
Db 840 CGAGCTGACTCTTGAGCCTTAACCAAGATGAGAGGAGAGAGATTTGGAGTTTGAG 899
QY 219 ThrLeuCysAlaValLeuGlnGluArgLysGlyGluLeuGlnAlaLeuAlaArgGlu 238

```

RESULT 10  
 US-09-925-301-101  
 ; Sequence 101, Application US/09925301  
 ; Patent No. US20020052308A1



```

Db 379 TTACCTGTTAATTTTGCACCTAGGCGCTATTATTGAAGACGACGACAGACCATCCA 320
QY 115 ProleuHISAlaLysAlaGluGlnHisLeuMetCysGluGlnHisGluSpGluysIle 134
Db 319 GATATTT-----GTCACTGCGCCCTGGAACATTTACAGGCAACATTA 281
QY 135 AsnIleTyrcysLeuSerCysGluValProThrCysSerLeuCysLysValPheGlyAla 154
Db 280 AATGTTTACTGTCTATTAGATAAATAATTAGTTGTGTCATTTGCTTACCATAGTCTCA 221
QY 155 HisLysAspCysGluValAlaProLeuProThrIleTyrcysArgGlnLysSerGluLeu 174
Db 220 CATATGCTCATCTATATAGATGACCTCAAAAGTGCCTATTGAAAGAAAAGACACTCCT 161
QY 175 SerAspGlyIleAlaMetLeuValAlaGlnAsnAspArgValGlnAlaValIleThrGln 194
Db 160 CAAAACGCTGTAACACTG-----ACTGACACACACTGACAGATCTTACCCAT 110
QY 195 MetGluGluValCysGlnThrIleGluAsnSerArgArg-----GlnLys 210
Db 109 CTTATTGAAAAGCTGAAGAACAAAATCTCTCTGAGAAAATGATCCAGGCGATTAG 50
QY 211 GlnLeuLeuAsnGlnArgPheGluThrLeuCysAlaValLeuGluGln 226
Db 49 GAAGCTGTCTCCAGTATTTTAAGAGCTTAATGATACATTAGAACAG 2

```

RESULT 13  
US-09-764-877-2549

; Sequence 2549, Application US/09764877  
; Patent No. US20020147140A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877

; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2549  
; LENGTH: 599

; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-764-877-2549

Alignment Scores:

```

Pred. No.: 5,23e-17 Length: 599
Score: 225.00 Matches: 59
Percent Similarity: 43.52% Conservative: 35
Best Local Similarity: 27.31% Mismatches: 98
Query Match: 11.77% Indels: 24
Gaps: 6

```

US-09-908-988b-2 (1-366) x US-09-764-877-2549 (1-599)

```

QY 17 MetAspAsnLeuGlnLysGlnLeuIleCysProIleCysLeuGluMetPheSerLysPro 36
Db 5 ATGCACAAATTTTGGAGAAAGTAACTTGCCATATGTTATAGATTATTTTGAAGATCCT 64
QY 37 ValValIleLeuProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGln 56
Db 65 CGGTGA---CTGCATCTCTCATACATTTTGTGGAATAATTGTTGGAAACATCTCTCAG 121
QY 57 AlaSerAsnPro-----LeuTrpGlnSerArgLysSerThrValSerSerGlyGly 74
Db 122 GCATCTGTAACCTTTTATATATGAGACCTTACGAATTCCA----- 163
QY 75 ArgPheArgCysProSerCysArgHisGlnValValLeuAspArgHisGlyValTyrcly 94
Db 164 ----CTCAAGTGCCTTAATTTGCAGAAAGTATTACTGAAATTTGCTCCAAAGCGCATTTGAATCT 220
QY 95 LeuGlnArgAsnLeuLeuValGluAsnIleIleAspIleTyrclyGlnLysSerSerArg 114

```

```

Db 221 TTACCTGTTAATTTTGCACCTAGGCGCTATTATTGAAGACGACGACAGACCATCCA 280
QY 115 ProleuHISAlaLysAlaGluGlnHisLeuMetCysGluGlnHisGluSpGluysIle 134
Db 281 GATATTT-----GTCACTGCGCCCTGGAACATTTACAGGCAACATTA 319
QY 135 AsnIleTyrcysLeuSerCysGluValProThrCysSerLeuCysLysValPheGlyAla 154
Db 320 AATGTTTACTGTCTATTAGATAAATAATTAGTTGTGTCATTTGCTTACCATAGTCTCA 379
QY 155 HisLysAspCysGluValAlaProLeuProThrIleTyrcysArgGlnLysSerGluLeu 174
Db 380 CATATGCTCATCTATATAGATGACCTCAAAAGTGCCTATTGAAAGAAAAGACACTCCT 439
QY 175 SerAspGlyIleAlaMetLeuValAlaGlnAsnAspArgValGlnAlaValIleThrGln 194
Db 440 CAAAACGCTGTAACACTG-----ACTGACACACACTGACAGATCTTACCCAT 490
QY 195 MetGluGluValCysGlnThrIleGluAsnSerArgArg-----GlnLys 210
Db 491 CTTATTGAAAAGCTGAAGAACAAAATCTCTCTGAGAAAATGATCCAGGCGATTAG 550
QY 211 GlnLeuLeuAsnGlnArgPheGluThrLeuCysAlaValLeuGluGln 226
Db 551 GAAGCTGTCTCCAGTATTTTAAGAGCTTAATGATACATTAGAACAG 598

```

RESULT 14  
US-09-764-877-2550

; Sequence 2550, Application US/09764877  
; Patent No. US20020147140A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877

; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2550  
; LENGTH: 599

; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-764-877-2550

Alignment Scores:

```

Pred. No.: 5,23e-17 Length: 599
Score: 225.00 Matches: 59
Percent Similarity: 43.52% Conservative: 35
Best Local Similarity: 27.31% Mismatches: 98
Query Match: 11.77% Indels: 24
Gaps: 6

```

US-09-908-988b-2 (1-366) x US-09-764-877-2550 (1-599)

```

QY 17 MetAspAsnLeuGlnLysGlnLeuIleCysProIleCysLeuGluMetPheSerLysPro 36
Db 5 ATGCACAAATTTTGGAGAAAGTAACTTGCCATATGTTATAGATTATTTTGAAGATCCT 64
QY 37 ValValIleLeuProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGln 56
Db 65 CGGTGA---CTGCATCTCTCATACATTTTGTGGAATAATTGTTGGAAACATCTCTCAG 121
QY 57 AlaSerAsnPro-----LeuTrpGlnSerArgLysSerThrValSerSerGlyGly 74
Db 122 GCATCTGTAACCTTTTATATATGAGACCTTACGAATTCCA----- 163
QY 75 ArgPheArgCysProSerCysArgHisGlnValValLeuAspArgHisGlyValTyrcly 94
Db 164 ----CTCAAGTGCCTTAATTTGCAGAAAGTATTACTGAAATTTGCTCCAAAGCGCATTTGAATCT 220
QY 95 LeuGlnArgAsnLeuLeuValGluAsnIleIleAspIleTyrclyGlnLysSerSerArg 114

```

```
DB 221 TTACCTGTATTTTGGACCTAAGGCTTTATTTGAAGTACACGACAAAGACATCCA 280
      ||| ||| : : ||||| : : |||||
      ||| ||| : : ||||| : : |||||
QY 115 ProleuHISAlaLysAlaGluGlnHisLeuMetCysGluGlnHisGluAspGluLysIle 134
      : : : : ||| ||||| : : : :
      : : : : ||| ||||| : : : :
DB 281 GATATT-----GTCACTGCCCTGAAACATTTACAGGCAACCATTA 319
      : : : : ||| ||||| : : : :
QY 135 AsnIleTyrcysLeuSerCysGluValProThrCysSerLeuCysLysValPheGlyAla 154
      ||| ||| ||| ||| : : ||| |||
      ||| ||| ||| ||| : : ||| |||
DB 320 AAGTTTACGTCTATTAGATGATAAAAAATTAAGTTTGTGTCATCTTACCATAGGTCAA 379
      ||| ||| ||| ||| : : ||| |||
QY 155 HisLysAspCysGluValAlaProLeuProThrIleTyrcysArgGlnLysSerGluLeu 174
      ||| : : : : ||| : : : : ||| : : : : |||
      ||| : : : : ||| : : : : ||| : : : : |||
DB 380 CATCATGTCATCTCTATAGATGACCTTCAAGTGCCTATTGAAAGAAAGACACTCCT 439
      : : : : ||| : : : : ||| : : : : |||
QY 175 SeraspGlyIleAlaMetLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGln 194
      : : : : ||| : : : : ||| : : : : |||
      : : : : ||| : : : : ||| : : : : |||
DB 440 CAAAAAGCTGTTGAACAGTTG-----ACTGACACACACTGGACAGATCTTACCCAT 490
      : : : : ||| : : : : ||| : : : : |||
QY 195 MetGluGluValCysGlnThrIleGlnAspAsnSerArgArg-----GlnLys 210
      : : : : ||| : : : : ||| : : : : |||
      : : : : ||| : : : : ||| : : : : |||
DB 491 CTATTGTAAAGCTGAAGAAACAAAATCTCTGAGAAAATGATCCAAAGCGATTAAG 550
      : : : : ||| : : : : ||| : : : : |||
QY 211 GlnLeuLeuAsnGlnArgPheGluThrLeuCysAlaValLeuGluGln 226
      : : : : ||| : : : : ||| : : : : |||
      : : : : ||| : : : : ||| : : : : |||
DB 551 GAAGCTGTCTCCAGATTTTAAGAGCTTATGATACATTGAAACAG 598
      : : : : ||| : : : : ||| : : : : |||

RESULT 15
US-09-764-864-46
; Sequence 46, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PFT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-46

Alignment Scores:
Pred. No.: 5.58e-17 Length: 625
Score: 225.00 Matches: 59
Percent Similarity: 43.52% Conservative: 35
Best Local Similarity: 27.31% Mismatches: 98
Query Match: 11.77% Indels: 24
DB: 10 Gaps: 6

US-09-908-988b-2 (1-366) x US-09-764-864-46 (1-625)
QY 17 MetaspAsnLeuGluLysGlnLeuIleCysProIleCysLeuGlnMetPheSerLysPro 36
      ||| ||| ||| ||| : : ||| ||| ||| |||
      ||| ||| ||| ||| : : ||| ||| ||| |||
DB 32 ATGCACAAATTTTGGAGAGAGTTGACTGTGCCATATGTTATATGTTATTTTGAAGATCCT 91
      : : : : ||| : : : : ||| : : : : |||
      : : : : ||| : : : : ||| : : : : |||
QY 37 ValValIleLeuProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGln 56
      ||| ||| ||| ||| ||| : : ||| ||| |||
      ||| ||| ||| ||| ||| : : ||| ||| |||
DB 92 CGGTGA---CTGCCATGCTCTCATACATTTTGTAGAAATTTGTTGGAACAACTTCTCAG 148
      : : : : ||| : : : : ||| : : : : |||
      : : : : ||| : : : : ||| : : : : |||
QY 57 AlaSerAsnPro-----LeuTrpGlnSerArgGlySerThrValSerSerGlyGly 74
      ||| ||| : : : : ||| : : : : |||
      ||| ||| : : : : ||| : : : : |||
DB 149 GCATCTGTACTTTTATATATGAGACCTTACGAATTC----- 190
      : : : : ||| : : : : ||| : : : : |||
QY 75 ArgPheArgCysProSerCysArgHisGluValAlaLeuAspArgHisGlyValIleTyrcly 94
      : : : : ||| : : : : ||| : : : : |||
      : : : : ||| : : : : ||| : : : : |||
DB 191 ---CTCAAGTGCCTTAATTTGCAAGATATTACTGAATTTGCTCCAACTGCATTTGAATCT 247
      : : : : ||| : : : : ||| : : : : |||
      : : : : ||| : : : : ||| : : : : |||
QY 95 LeuGlnArgAsnLeuLeuValGluAsnIleIleAspIleTyrcLysGlnGlnSerSerArg 114
```

```
DB 248 TTACCTGTATTTTGGACCTAAGGCTTATTTGAAGTACACGACAAAGACATCCA 307
      ||| ||| : : ||||| : : |||||
      ||| ||| : : ||||| : : |||||
QY 115 ProleuHISAlaLysAlaGluGlnHisLeuMetCysGluGlnHisGluAspGluLysIle 134
      : : : : ||| ||||| : : : :
      : : : : ||| ||||| : : : :
DB 308 GATATT-----GTCACTGCCCTGAAACATTTACAGGCAACCATTA 346
      : : : : ||| ||||| : : : :
QY 135 AsnIleTyrcysLeuSerCysGluValProThrCysSerLeuCysLysValPheGlyAla 154
      ||| ||| ||| ||| : : ||| |||
      ||| ||| ||| ||| : : ||| |||
DB 347 AAGTTTACGTCTATTAGATGATAAAAAATTAAGTTTGTGTCATCTTACCATAGGTCAA 406
      ||| ||| ||| ||| : : ||| |||
QY 155 HisLysAspCysGluValAlaProLeuProThrIleTyrcysArgGlnLysSerGluLeu 174
      ||| : : : : ||| : : : : ||| : : : : |||
      ||| : : : : ||| : : : : ||| : : : : |||
DB 407 CATCATGTCATCTCTATAGATGACCTTCAAGTGCCTATTGAAAGAAAGACACTCCT 466
      : : : : ||| : : : : ||| : : : : |||
QY 175 SeraspGlyIleAlaMetLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGln 194
      : : : : ||| : : : : ||| : : : : |||
      : : : : ||| : : : : ||| : : : : |||
DB 467 CAAAAAGCTGTTGAACAGTTG-----ACTGACACACACTGGACAGATCTTACCCAT 517
      : : : : ||| : : : : ||| : : : : |||
QY 195 MetGluGluValCysGlnThrIleGlnAspAsnSerArgArg-----GlnLys 210
      : : : : ||| : : : : ||| : : : : |||
      : : : : ||| : : : : ||| : : : : |||
DB 518 CTATTGTAAAGCTGAAGAAACAAAATCTCTGAGAAAATGATCCAAAGCGATTAAG 577
      : : : : ||| : : : : ||| : : : : |||
QY 211 GlnLeuLeuAsnGlnArgPheGluThrLeuCysAlaValLeuGluGln 226
      : : : : ||| : : : : ||| : : : : |||
      : : : : ||| : : : : ||| : : : : |||
DB 578 GAAGCTGTCTCCAGATTTTAAGAGCTTATGATACATTGAAACAG 625
      : : : : ||| : : : : ||| : : : : |||

RESULT 16
US-09-764-877-269
; Sequence 269, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 269
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-269

Alignment Scores:
Pred. No.: 5.58e-17 Length: 625
Score: 225.00 Matches: 59
Percent Similarity: 43.52% Conservative: 35
Best Local Similarity: 27.31% Mismatches: 98
Query Match: 11.77% Indels: 24
DB: 10 Gaps: 6

US-09-908-988b-2 (1-366) x US-09-764-877-269 (1-625)
QY 17 MetaspAsnLeuGluLysGlnLeuIleCysProIleCysLeuGlnMetPheSerLysPro 36
      ||| ||| ||| ||| : : ||| ||| ||| |||
      ||| ||| ||| ||| : : ||| ||| ||| |||
DB 32 ATGCACAAATTTTGGAGAGAGTTGACTGTGCCATATGTTATATGTTATTTTGAAGATCCT 91
      : : : : ||| : : : : ||| : : : : |||
      : : : : ||| : : : : ||| : : : : |||
QY 37 ValValIleLeuProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGln 56
      ||| ||| ||| ||| ||| : : ||| ||| |||
      ||| ||| ||| ||| ||| : : ||| ||| |||
DB 92 CGGTGA---CTGCCATGCTCTCATACATTTTGTAGAAATTTGTTGGAACAACTTCTCAG 148
      : : : : ||| : : : : ||| : : : : |||
      : : : : ||| : : : : ||| : : : : |||
QY 57 AlaSerAsnPro-----LeuTrpGlnSerArgGlySerThrValSerSerGlyGly 74
      ||| ||| : : : : ||| : : : : |||
      ||| ||| : : : : ||| : : : : |||
DB 149 GCATCTGTACTTTTATATATGAGACCTTACGAATTC----- 190
      : : : : ||| : : : : ||| : : : : |||
QY 75 ArgPheArgCysProSerCysArgHisGluValAlaLeuAspArgHisGlyValIleTyrcly 94
      : : : : ||| : : : : ||| : : : : |||
      : : : : ||| : : : : ||| : : : : |||
DB 191 ---CTCAAGTGCCTTAATTTGCAAGATATTACTGAATTTGCTCCAACTGCATTTGAATCT 247
      : : : : ||| : : : : ||| : : : : |||
      : : : : ||| : : : : ||| : : : : |||
QY 95 LeuGlnArgAsnLeuLeuValGluAsnIleIleAspIleTyrcLysGlnGlnSerSerArg 114
```

Db 248 TTACCTGTTAATTTTGGACATAAGGCTATTATTGAAGAACACAGACAGACATCCA 307  
 Oy 115 ProeuHisAlaValuLglnHisLeuMetCysGluHisGlnuSpgluysile 134  
 Db 308 GATATT-----GTACCTGCGCCCTGAACATATTACGGCAACCATTA 346  
 Oy 135 AsnIleTyrCysLeuSerCysGluValProThrCysSerLeuCysLysValPheGlyAla 154  
 Db 347 AATGTTTACTGTCTATAGATAAAATATTAGTTTGTGTCATTCGCTTACCATAGTCA 406  
 Oy 155 HisLysAspCysGluValAlaProleuProThrIleTyrLysArgGlnLysSerGluLeu 174  
 Db 407 CATATAGGTCACTCTATAGATGACCTTCAAGTGCCTATTGAAAGAAAGACACTCCT 466  
 Oy 175 SerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGln 194  
 Db 467 CAATAACGTCTGAACACTG-----ACTGCACACACTGCAGACATCTTACCCAT 517  
 Oy 195 MetGluGluValCysGlnThrIleGlnuSpgAsnSerArgArg-----GlnLys 210  
 Db 518 CTATTGTGAAGAGCTGAAGAACAAATCTCATCTCTGAGAAATGATCCAGGCCATTAG 577  
 Oy 211 GlnLeuEuAsnGlnArgPheGlnThrLeuCysAlaValLeuGluGlu 226  
 Db 578 GAACCTGTTCTCAGTATTTTAAGAGCTTAATGATACATTAGAACAG 625

## RESULT 17

US-09-728-952-1

; Sequence 1, Application US/09728952  
 ; Patent No. US20020111302A1

## GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Zhou, Ping  
 ; APPLICANT: Goodrich, Ryle  
 ; APPLICANT: Liu, Chenghua  
 ; APPLICANT: Asundi, Vinod  
 ; APPLICANT: Wang, Jian-Rui  
 ; APPLICANT: Yamazaki, Yuki  
 ; APPLICANT: Ujwal, Manusha L.  
 ; APPLICANT: Drmanac, Radoje T.  
 ; TITLE OF INVENTION: No. US20020111302A1el Nucleic Acids and  
 ; FILE REFERENCE: 799  
 ; CURRENT APPLICATION NUMBER: US/09/728,952  
 ; CURRENT FILING DATE: 2000-11-30  
 ; NUMBER OF SEQ. ID NOS: 101  
 ; SOFTWARE: PC\_FL\_genes Version 2.0  
 ; SEQ. ID NO 1  
 ; LENGTH: 7596  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(7596)  
 ; US-09-728-952-1

## Alignment Scores:

Pred. No.: 3,65e-15 Length: 7596  
 Score: 223.50 Matches: 85  
 Percent Similarity: 39.07% Conservative: 67  
 Best Local Similarity: 21.85% Mismatches: 165  
 Query Match: 11.69% Indels: 72  
 DB: 10 Gaps: 13

US-09-908-988b-2 (1-366) x US-09-728-952-1 (1-7596)

Oy 14 AlaHisAsnMetAspAsnLeuGluLysGlnLeuIleCys-ProIleCys--LeuGluMe 32  
 Db 2805 GCTCAT-----CATCTCCCGCGGCGCAAAATCTGCTCTGCTGCTCTAGTTT 2855  
 Oy 32 tPheserLysProValValIleLeuProCysGlnHisAsnLeuCysArgLysCysAlaAs 52

Db 2856 CACCATCAGACAGCTCATTA-----TGTCTCCA 2882  
 Oy 52 nAspValPheGlnAlaSerAsnProLeuTPrGlnSerArgLysSerThrValSerSe 72  
 Db 2883 AATATATATCCCTCCAG-----ACCTGACACCTTATCC-- 2916  
 Oy 72 rGlyLysArgPheArgCysProSerCysArgHisGluValIleLeuAspArgHisGlyVa 92  
 Db 2917 -----TGTCCAGTATGCGCGGCAGACGTCATCTCCACAGACAGGGGCT 2960  
 Oy 92 lTyrGlyLeuGlnArgAsnLeuValGluAsnIleLeuAspIleTyrLysGln----- 110  
 Db 2961 CTCGGCACTGCAGAAACAATCTTTCATCAGACGCTCATGGAGGCAATGACAGACGACC 3020  
 Oy 111 -----GluSerSerArgProLeuHisAlaLysAlaGlnHisLe 124  
 Db 3021 TGATGGGGCCACAGACCCCGGAGACCCGCCCTCAGTGTAGTGGCTGGCCCTCT 3080  
 Oy 124 uMetCysGluGlnHisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValPr 144  
 Db 3081 CTCCTGCCCAACCATATAGGCAAGACGATGAGTTTACTGTGAGGCTGTGAGACGCGC 3140  
 Oy 144 oThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGluValAlaProLeuPr 164  
 Db 3141 CATGTGTGTAGTGGCGCGCC---GGGAGACATCGTATGAGCATGGCACAGTGTCTGAG 3197  
 Oy 164 oThrIleTyrLysArgGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGl 184  
 Db 3198 GATGTGTGTAGTGGCGCGCCAGACAGCGCCCTCAGCGCTGAGGCTGGCGGCGC 3257  
 Oy 184 yAsnAspArgValGlnAlaValIleThrGlnMetGluGluValCysGlnThrIleGluAs 204  
 Db 3258 ATTGCCACAGCTGTCCGACAGAAATTCCTTATGCGGGGCATCCAGCCAGCTGACGA 3317  
 Oy 204 pAsnSerArgArgGlnLysGlnLeuEuAsnGlnArgPheGlnThrLeuCysAlaValIle 224  
 Db 3318 GCCCAAGCAGAGCCCTGCGCCAGATCATGTCAGCGCTTCAGGACCTGGAGCAAGCACT 3377  
 Oy 224 uGluGluArgLysGlnLysLeu-----LeuGlnAlaLeuAlaArgGlnGlnGlu 241  
 Db 3378 GCACAGCGCAAGCAGCTGTGTCAAGCAGCACTGTGAGCCATTTGTGGGGCCAAACAA 3437  
 Oy 241 uLysLeuGlnArgValArgLysLeuIleArgGlnTyrGlyAspHisLeuGluYerSe 261  
 Db 3438 GGTGTTCAAAGCCAGCTGTGACACACTGCGCCAGGGTCAGGACATCGGCACTAGCTG 3497  
 Oy 261 rLysLeuValGlnSerAlaIleGlnSerMetGluGluProGlnMetAlaLeuTyrLeuG 281  
 Db 3498 CAGCTTTCAGAGCAGCAGCAGCTGCGCTGGGCTCGGCCCGGAGGTTGTGTGGCGCA 3557  
 Oy 281 nGlnAlaLysGlnLeuIleAsnLysValGlyAlaMetSerLysValGluLeuAlaGlyAr 301  
 Db 3558 GCACATGCGAGAG-----CGGCTGGCTGATTTGGCGGCACAGGCCCTTCCGGAGCG 3608  
 Oy 301 qProGluProGlyTyrGlu-----SerMetGluGlnPheSerValSerVa 316  
 Db 3609 GCCACATGAGATATGACATGCACTGCACTGCTTGTAGGTGACGCTGTGCGGCATCGT 3668  
 Oy 316 lGluHisValAlaGluMetLeuArgThr----- 325  
 Db 3669 GCTCAATCTGGCGCACTGCTCACACAGAGGCCACATGACAGCAAAACGTTGGCCACGGG 3728  
 Oy 326 -----lLeasPheGlnProGlyAlaAlaGlyAspGlnGluAspAs 339  
 Db 3729 AGAGGGCTCGCCAGCGGCTAGTGGGCGCAGCTGCTGCTCACTGTCACTACCAAAAG 3788  
 Oy 339 pAspMetAlaLeuAspGlyLysGlnGluGlyAsnAlaGlyLeuGlnGluGluArgLeuAspVa 359  
 Db 3789 CAGG-----GACGGGGGTTGTGTGTGGCCACAGGACGCGCTGACGCTGACAGAT 3839  
 Oy 359 l-----ProGluGlySerGlyLeu 365









GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OW protein - nucleic search, using frame\_plus\_p2n model

Run on: December 3, 2002, 12:14:13 ; Search time 56 Seconds

(without alignments) 2004.352 Million cell updates/sec

Title: US-09-908-988B-2

Perfect score: 1912

Sequence: 1 MNFTVGFKPLGDHNMNDL.....EGNAGLEERLDVPEGSLH 366

Scoring table:

BLOSUM62

Xgapop 10.0, Xgapext 0.5

Xgapop 6.0, Xgapext 7.0

Delop 6.0, Delext 7.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framed-p2n.model -DEV=x1h

-O=/cg2\_1/USPTO.spool/US09908888/runat\_26112002\_112154\_11213/app.query.fasta\_1.519

-DB=issued.patents.na -OPMT=fastp -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi

-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=20

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USRR=US09908888.ecgn.1.1.17.ecgnat\_26112002\_112154\_11213 -NCPU=6 -ICPU=3

-NO.XLPHY -NO.MAP -IAREQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=7

-WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued.patents.na.\*

1: /cg2\_1/ptodata/1/ina/5A.COMB.seq.\*

2: /cg2\_6/ptodata/1/ina/5B.COMB.seq.\*

3: /cg2\_6/ptodata/1/ina/6A.COMB.seq.\*

4: /cg2\_6/ptodata/1/ina/6B.COMB.seq.\*

5: /cg2\_6/ptodata/1/ina/PCNUS.COMB.seq.\*

6: /cg2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1101.5	57.6	1757	4	US-09-484-970B-37
2	255.5	13.4	3262	4	US-09-561-989-9
3	201	10.5	2854	2	US-08-724-394A-17
4	178	9.3	3446	4	US-09-484-970B-93
5	166.5	8.7	3312	1	US-08-049-473-1
6	166.5	8.7	3312	1	US-08-312-648-1
7	166.5	8.7	3312	5	PCT-US94-04190-1
8	132	6.9	2155	2	US-08-095-728B-5
9	132	6.9	2155	5	PCT-US92-02320A-5
10	132	6.9	3511	3	US-08-892-747-13
11	131	6.9	246240	2	US-08-724-394A-20
12	131	6.9	246240	2	US-08-724-394A-21

13	131	6.9	246240	2	US-08-724-394A-22	Sequence 22, Appl
14	127	6.6	1767	1	US-07-903-466-2	Sequence 2, Appl
15	127	6.6	1767	5	PCT-US93-05794-2	Sequence 2, Appl
16	127	6.6	3018	1	US-07-903-466-1	Sequence 1, Appl
17	127	6.6	3018	5	PCT-US93-05794-1	Sequence 1, Appl
18	123	6.4	5173	4	US-08-801-308-2	Sequence 2, Appl
19	122.5	6.4	2240	4	US-08-697-610-1	Sequence 1, Appl
20	122.5	6.4	2240	4	US-08-349-357-1	Sequence 1, Appl
21	122.5	6.4	2301	4	US-09-085-199B-8	Sequence 8, Appl
22	122	6.4	3036	2	US-08-306-691B-52	Sequence 52, Appl
23	122	6.4	3036	5	PCT-US92-02320A-1	Sequence 1, Appl
24	122	6.4	3036	5	PCT-US92-02320A-1	Sequence 1, Appl
25	118.5	5.9	2262	4	US-08-639-237-1	Sequence 2, Appl
26	113.5	5.9	2248	1	US-08-975-405-1	Sequence 1, Appl
27	113.5	5.9	2248	4	US-09-167-109-6	Sequence 6, Appl
28	113.5	5.9	2264	4	US-08-897-340-6	Sequence 6, Appl
29	112.5	5.9	2121	2	US-09-252-329-6	Sequence 6, Appl
30	112.5	5.9	2121	3	PCT-US95-16980-1	Sequence 1, Appl
31	112.5	5.9	2359	5	US-08-074-879-9	Sequence 9, Appl
32	112	5.9	7833	1	US-08-468-057A-9	Sequence 9, Appl
33	112	5.9	7833	1	US-08-468-057A-9	Sequence 9, Appl
34	111.5	5.8	2455	4	US-09-167-109-3	Sequence 3, Appl
35	111	5.8	4722	4	US-08-979-608A-14	Sequence 14, Appl
36	111	5.8	5707	1	US-08-425-061-11	Sequence 11, Appl
37	111	5.8	5707	1	US-08-825-886-11	Sequence 11, Appl
38	111	5.8	5709	1	US-08-425-061-7	Sequence 7, Appl
39	111	5.8	5709	1	US-08-425-061-8	Sequence 8, Appl
40	111	5.8	5709	1	US-08-425-061-9	Sequence 9, Appl
41	111	5.8	5709	1	US-08-825-886-7	Sequence 7, Appl
42	111	5.8	5709	1	US-08-825-886-8	Sequence 8, Appl
43	111	5.8	5709	1	US-08-825-886-9	Sequence 9, Appl
44	111	5.8	5710	1	US-08-425-061-6	Sequence 6, Appl
45	111	5.8	5710	1	US-08-825-886-6	Sequence 6, Appl

## ALIGNMENTS

RESULT 1

US-09-484-970B-37

; Sequence 37, Application US/09484970B

; Patent No. 6426186

; GENERAL INFORMATION:

; APPLICANT: Jones, Karen A.

; APPLICANT: Volkman, Wayne

; APPLICANT: Walker, Michael G.

; TITLE OF INVENTION: BONE REMODELING GENES

; FILE REFERENCE: PB-0014 US

; CURRENT APPLICATION NUMBER: US/09/484,970B

; CURRENT FILING DATE: 2000-01-18

; NUMBER OF SEQ ID NOS: 172

; SOFTWARE: PERL Program

; SEQ ID NO 37

; LENGTH: 1757

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; OTHER INFORMATION: Incyte ID No. 6426186 247384.1CB1

US-09-484-970B-37

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-908-988B-2 (1-366) x US-09-484-970B-37 (1-1757)

4.9e-118  
1101.50  
78.29%  
60.86%  
57.61%  
4  
4  
1757  
213  
61  
11

QY 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluysGlnLeuIleuIleCysProIleCys 29  
DB 143 CTGATCCAGATGGGATATCCCATGAGAGAACTTGAGAGAAAGCAAGCTGATCTGCTATCTGC 202

QY LeuGlumetPheSerLySProValIleLeuProCysGlnHisAsnLeuCyArgLys 49  
 |||||  
 Db CTGAGATGTTTACCAAGCCAGTGTCTATCTCCGTCGACACACACTGTCCCGAAG 262  
 QY CysAlaAsnAspValPheGlnAlaSerAsnProLeuTyrPglInSerArgLysThr 69  
 |||||  
 Db TGTGCCAATGATATTTTCCAGGCGCTTAACCCGTATTTCGCCCAAGAGAGTACCACC 322  
 QY ValSerSerGlyArgPheArgCysProSerCysArgHisGlnValIleLeuAspArg 89  
 |||||  
 Db ATGGCATCAGGGGCGCATTCGCGCTGCGCATCTGTAGACATGATGGTTTGGATAGA 382  
 QY HisGlyValTyrGlyLeuGlnArgAsnLeuValGluAsnIleLeuAspIleTyrLys 109  
 |||||  
 Db CATGGGTATATGACTTCAAGAGAACTGCTGGTGGAGAACATCATCTGACATCTACAAA 442  
 QY GlnGlnLys--SerSerArgProLeuHisAlaLysAlaGlnGlnHisLeuMetCysGln 128  
 |||||  
 Db CAGAGAGCTTCACGTCCGCGCTG-----CAGAGGGCAGTCACCCCATGTGTCAGAGAG 496  
 QY HisGlnAspGlnLysIleAsnIleTyrCysLeuSerCysGlnValProThrCysSerLeu 148  
 |||||  
 Db CACGAAATGAGAAATCAACATCTACTGTCTACGTGTGAGGTGCGCACCTCTCCATG 556  
 QY CysLysValPheGlyAlaHisLysAspCysGlnValAlaProLeuProThrIleTyrLys 168  
 |||||  
 Db TGCAGAGTGTGGGATCCACAGAGCGCTGCGAGTGGCCCATGTGACAGTGTCTTCAG 616  
 QY ArgGlnLysSerGlnLeuSerAspGlyIleAlaMetLeuValAlaGlnAsnAspArgVal 188  
 |||||  
 Db GGCACAAAGATGAACTGAATAACTGATCTCCATGCTGTGCGGGGAAATGACCGGTG 676  
 QY GlnAlaValIleThrGlnMetGlnGlnValCysGlnThrIleGlnAspAsnSerArgArg 208  
 |||||  
 Db CAGGCCATCATCTACCTGAGTGGAGATCCCTCGAGTCCAGCAAGGAGGACAGTCCAG 736  
 QY GlnLysLeuLeuAsnGlnArgPheGlnThrLeuCysAlaValLeuGlnArgLys 228  
 |||||  
 Db GTTAAAGAGAGCTGAGCCAGAGTTTGCACGTGTGATGCCATCTCGATGAGAAAGAA 796  
 QY GlnGlnLeuLeuGlnAlaLeuAlaArgGlnGlnGlnLysLeuGlnArgValArgGly 248  
 |||||  
 Db AGTGAAGTGTGCTGACGCGATGACGAGGAGCAGGAAAGGCTTTCGCTTCAGAGGCC 856  
 QY LeuIleArgGlnTyrGlyAspHisLeuGlnGlySerSerLysLeuValGlnSerAlaIle 268  
 |||||  
 Db CTCATCCAGCAGTACAGAGAGAGCTGACAGTCCACAAAGCTGTGTGGAAGTCCATC 916  
 QY GlnSerMetGlnGlnProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGlnLeuIleAsn 288  
 |||||  
 Db CAGTCCCGGAGAGAGCTGGGGAGCCACCTTCCTGACTGCCAAGCAATCATCAAA 976  
 QY LysValGlnAlaMetSerLysValGlnLeuAlaGlyArgProGlnProGlyTyrGlnLys 308  
 |||||  
 Db AGCATTGTGAGAGCTTCCAGGGCTGCCAGCTGGGGAGACAGAGAGGCGCTTGAAGAAC 1036  
 QY MetGlnGlnPheSerValSerValGlnHisValAlaGlnMetLeuArgThrIleAspPhe 328  
 |||||  
 Db ATGACACTTCTTACTTGTGATTAGACATAGCAGAGAGCCCTGAGAGGCATTTGACTTT 1096  
 QY GlnProGlnAlaAlaGlyAspGlnGlnAspAspPheMetAla-----LeuAsp 344  
 |||||  
 Db 1097 -----GGGACAGATGAGAGAGAGAGAAATTCATTGAGAGAGAAATCATGAGAA 1144  
 QY 345 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 354  
 |||||  
 Db 1145 GAGGAGAGTCCACAG 1174  
 |||||

; GENERAL INFORMATION:  
 ; APPLICANT: KOLLER, Klaus-Peter  
 ; TITLE OF INVENTION: No. 6468750e1 Cell Regulation Factor T1020  
 ; FILE REFERENCE: seq  
 ; CURRENT APPLICATION NUMBER: US/09/561, 989  
 ; CURRENT FILING DATE: 2000-04-27  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 9  
 ; LENGTH: 3262  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-561-989-9  
 Alignment Scores:  
 Pred. No.: 2,46e-19 Length: 3262  
 Score: 255.50 Matches: 67  
 Percent Similarity: 41.738 Conservative: 44  
 Best Local Similarity: 25.198 Mismatches: 102  
 Query Match: 13.36% Indels: 53  
 DB: Gaps: 2  
 US-09-908-988b-2 (1-366) x US-09-561-989-9 (1-3262)  
 QY 78 CysProSerCysArgHisGlnValIleLeuAspArgHisGlnValTyrGlyLeuGlnArg 97  
 |||||  
 Db 37 TGCACCAATGATGACCGCAGCGCTCATCTGTGATGACCGGGGCGCTCCGCTCCCAAG 96  
 QY 98 AsnLeuValGlnLysIleLeuAspIleTyrGlnGlnLysSerArgProLeu--- 116  
 |||||  
 Db 97 AATCGCGTGTGAAAGGGGTAATTGACCGCTTCCACAGAGAGCAAAAGCGCGCCCTCAAG 156  
 QY 117 -----HisAlaLysAlaGlnGlnHisLeuMetCysGln----- 127  
 |||||  
 Db 157 TCCCAAGCTCGGAGAGAGGGCCCAAGAGAGCCAGCTCATGTGCGAAGATGGATGTC 216  
 QY 127 ----- 127  
 |||||  
 Db 217 TTTCTACTGCGATCCGTGCGCGCTGCGCTGCCACCGCGCGGGGGCCCTAGCCAGAC 276  
 QY 127 ----- 127  
 |||||  
 Db 277 CCGCTGTGCTCCCGCCCGCCAGGCTCGTGTACCGCGAGCTGAGCCAGCCAGAGTCTCC 336  
 QY 128 -----GlnHisGlnLysGlnLysIleAsnIleTyrCysLeuSerCysGlnValPro 144  
 |||||  
 Db 337 ACGTCACAGACAGCAGAGCTGAGAACACACAGCATGTACTGCGATGCAAGATGCC 396  
 QY 145 ThrCysSerLeuCysLysValIlePheGlyAlaHisLysAspCysGlnValAlaProLeuPro 164  
 |||||  
 Db 397 GTGTGCTTACCAAGTGTGGAGAGGGCCAAACACTCCAGCCAGCAAGTCAAGCTTGGGG 456  
 QY 165 ThrIleTyrLysArgGlnLysSerGlnLeuSerAspGlyIleAlaMetLeuValAlaGly 184  
 |||||  
 Db 457 GCCATGTGAGAACTACATAGAGCCACTCTCCAGGCGGTGAAGCGACTGCACAGAG 516  
 QY 185 AsnAspArgValGlnAlaValIleThrGlnMetCysGlnValCysGlnThrIleGlnAsp 204  
 |||||  
 Db 517 GCCAAAGAGCCCAAGAGAGTTCGTGTGACAGTCCGCAACAGTGTCCACAGATCCAGAG 576  
 QY 205 AsnSerArgArgGlnLysGlnLeuLeuAsnGlnIleArgPheGlnThrIleCysAlaValLeu 224  
 |||||  
 Db 577 AACAGTGTGAGATTGAGAGCTGTGTGTGGGCCCAATGTGATGCCCTATGATGCCCTC 636  
 QY 225 GlnGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 244  
 |||||  
 Db 637 AACAGAGAGAAAGCCAGCGTGTGGCCCGGTCAACAGAGAGCATGACACAGAGCTGAAG 696  
 QY 245 ArgValArgGlnLeuIleArgGlnTyrGlyAspHisLeuGlnGlySerSerLysLeuVal 264  
 |||||  
 Db 697 GTGTGTCAGATCAGATCTCTCACTGACAGTGAATTTGCCACAGACACAGAGTCTCATG 756  
 QY 265 GlnSerAlaIleGlnSerMetGlnGlnProGlnMetAlaLeuTyrLeuGlnGlnAlaLys 284

DB 757 GAGTACTGCTGAGGTGATTAGGAAATGATCCTAGTGTGTTTTCAGATTCTGAC 816  
OY 285 GtuleuileasnllyVal 290  
DB 817 GCCCTCATTAAGAGAGTG 834

RESULT 3  
US-08-724-394A-17  
; Sequence 17, Application US/08724394A  
; Patent No. 5872237  
; GENERAL INFORMATION:  
; APPLICANT: Feder, John N.  
; APPLICANT: Krommal, Gregory S.  
; APPLICANT: Lauer, Peter M.  
; APPLICANT: Ruddy, David A.  
; APPLICANT: Thomas, Winston  
; APPLICANT: Tsuchihashi, Zenta  
; APPLICANT: Wolff, Roger K.  
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1  
; TITLE OF INVENTION: Sequences and Antibodies Thereo  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,394A  
; FILING DATE: 01-Oct-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitts, Renee A.  
; REGISTRATION NUMBER: 35,136  
; REFERENCE/DOCKET NUMBER: 017957-000100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-576-0200  
; TELEFAX: 415-576-0300  
; INFORMATION FOR SEQ. ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2854 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..2854  
; OTHER INFORMATION: /note= "cDNA 27"  
US-08-724-394A-17

Alignment Scores:  
Pred. No.: 4.25e-13 Length: 2854  
Score: 201.00 Matches: 90  
Percent Similarity: 38.12% Conservative: 56  
Best Local Similarity: 23.50% Mismatches: 161  
Query Match: 10.51% Indels: 76  
DB: 2 Gaps: 14

US-09-908-988b-2 (1-366) x US-08-724-394A-17 (1-2854)  
OY 22 LysGlnIleuileucysProileucysleuGluMetPheSerIysProvalIleleuPro 41  
DB 55 GAGGAAGCCACGCTGCTCATCTGCTGAGCTGATGAGAACCCAGTAGAGCATC---AAC 111  
OY 42 CysGlnHisAsnLeucysArgLysCysAlaAsnAspValPheGlnAlaSerAsnProleu 61

DB 112 TGTGACACAGCTACTGACCATCTGTGTATACAGACTTCTTTAA-----AACCCAGC 165  
OY 62 TrpGlnSerArgGlySerThrThrValSerSerClyGlyArgPheArgCysProSerCys 81  
DB 166 CAAGAAGCAACTAGAGCAGAGACA-----TTCTGCTGTCCCACTG 207  
OY 82 ArgHisGluValIleuAspArgHisGlyValIleuArgGlyLeuGlnArgAsnLeuVal 101  
DB 208 CGGGCTCCATTTTCATATGAT-----AGCTCCGACCCACACAGCAGCTG 252  
OY 102 GluAsnIleIleAspIleTyrIysGlnGluSerSerArgProleuHisAlaLysAlaGlu 121  
DB 253 GGAAGCCTCATTAAGACCCCTCAAGAG-----ACGAT 285  
OY 122 GlnHisLeuMetCysGlnGluHisGluAspGluLysIleAsnIleTyrCysLeuSerCys 141  
DB 286 CAGAAATGTCAATGTGAGAACAC---GGAGACAGCTTCACCTGTTCTCGAAGACGAG 342  
OY 142 GluValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGluValAla 161  
DB 343 GGGCAGCTCATCTGCTGGGGCTGTGAGCGGCACACAGCACAAAGGCACACACAGCT 402  
OY 162 ProleuProThrIleTyrIysArgGlnLysSerGluLeuSerAspGlyIleAlaMetLeu 181  
DB 403 CTGTGTGACAGAGTATGCCAGGCTACAGGAAGAAAGCTCAGAAAGCTGTGACAAACTG 462  
OY 182 ValAlaGlyAsnAspArg-----ValGlnAlaValIleThrGln 194  
DB 463 AAGCAACTGGAAGACAGATGTACGAGCAGAGCAGAGCTGTCCACACCATGCAGATTAAC 522  
OY 195 MetGluGluValCysGlnThrIleGluAspAsnSerArgArgGlnLysGlnLeuAsn 214  
DB 523 TGGAAAGAAAGGTACAGATT-----CAGACACAAAAAATCCG 561  
OY 215 GlnArgPheGlnThrLeuCysAlaValLeuGluGluArgLysGlyGluLeuGlnAla 234  
DB 562 TCTGACTTTAAGAACTCCAGTGTTCCTACATGAGAAAGAAAGTCTTCTCTGAGG 621  
OY 235 LeuAlaArgGluGlnGluLysLeuGlnArgValArgLysLeuIleGlnIleGly 254  
DB 622 CTGGAGAAAGAAACACACAGACTGTGAGTGA-----CTGAGGAGACTAGAG 669  
OY 255 AspHisLeuGluGlySerSerLysLeuValGluSerAlaIleGlnSerMetGluGluPro 274  
DB 670 GCTGCTCTGGGCTGAAGGATGAACTCAAGACCCACATCTCGGAACGAGAGAAA 729  
OY 275 -----GlnMetAlaLeuTyrLeuGlnGlnAlaLysGluLeuIleAsnLysValGly 291  
DB 730 TGTCAAGGCTCAGCCCAAGAAATGCTGACAGAAATGCAATGACACTTTCAGACAGATTGG 789  
OY 292 AlaMetSerLysValGluLeuAlaGlyArgProGluProGlyTyrGluSerMetGluGln 311  
DB 790 GCTGTG---AAGCTGGAACATCAGAGGCTGTCTCTGGAATTCATATCATGTGCAT 846  
OY 312 PheSerValSerValGlnHisValAlaGluMetLeuArg-----ThrIleAsp 327  
DB 847 GTTTCAGCTTACTTCGATGATGAGAAATGTTAAGAGTCAATCAAGTTAGTGTACT 906  
OY 328 PheGlnProGlyAlaAlaGlyAspGlu-----GluAspAspMetAla 336  
DB 907 CTGGATCCAGATACAGCTCTCAGCAAGTAATTCCTGTGAGATCCGAGACAGAGTACT 966  
OY 337 -----GluAspAspMetAla----- 342  
DB 967 CGTGATTAACCCAGGAGGATTCAGACACATCTTCCAGGAGATTTACTGCCCTCCCTCT 1026  
OY 343 ---LeuAspGlyGluGluGlyAsnAlaGlyLeuGluGluGluArgLeuAspValProGlu 361  
DB 1027 GCTTGTGGTTGTGAAGGCTTCACCTCAGGAAGAGCTTACTTGAAGAGGATGTTGGGAA 1086  
OY 362 GlySerGly 364









Db 901 CGTCAGAGAAATGGCTCTAAGTGTGTTGATGCTCATGTCGTGAAAAATTGATTTGG 960

[illegible]



Percent Similarity: 33.17% Conservative: 51  
 Best Local Similarity: 20.91% Mismatches: 154  
 Query Match: 6.90% Indels: 124  
 DB: 5 Gaps: 17

US-09-908-988b-2 (1-366) x PCT-US92-02320A-5 (1-2155)

QY 24 LeuileCysProIleCysLeuGluMetPheSerLysProValAlaIleuProCysGln 43  
 DB 243 CTGCGCTGCGCAGCAATGC---CAGCGCGAAGCCCAAGTCCCGAAGCTGCTGCTGCTG 299  
 QY 44 HisAsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnProLeuTyrGln 63  
 DB 300 CACACGCTGTGCTGCGATGCTGAG---GAGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 326  
 QY 64 SerArgLysSerThrThrValSerSerGlyLysArgPheArgCysProSerCysArg--- 82  
 DB 327 -----GCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 359  
 QY 83 -----HisGluValAlaLeuAspArgHisGlyValTyrGlyLeuGln 96  
 DB 360 CCGTGGCCCTAGTGCAGACACACCGCGCTGATACGCTCTTTTCGAGAGTGTGACG 419  
 QY 97 ArgAsnLeuValGlnAlaAsnIleIle----- 105  
 DB 420 CGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 479  
 QY 106 -----AspIleTyrLysGlnLysSerSerArgProLeuHisAlaLysAla---Glu 121  
 DB 480 GAGTGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539  
 QY 122 GlnHisLeuMetCysGlnGlnHisGlnAspGlyLysIle----- 134  
 DB 540 GCACACACAGTGTCTTCAACGACGAGCGCGCGCTGACAGAGCTGCGCAACAGTCG 599  
 QY 134 ----- 134  
 DB 600 GTGCGTAGTCTCTGAGCGGACCGCGCAACACATCTTCTGCTCCAAACCCCAAC 659  
 QY 135 -----AsnIleTyrCysLeuSerCysGluValProThrCysSer 147  
 DB 660 CACCGCACCCCTACGCTGACGACGATCTAGTCCGAGATGTTCCAGCGGCTGCTGCTG 719  
 QY 148 LeuCysLysValPhe---GlyAlaHisLysAsp---CysGluValAlaProLeuPro 164  
 DB 720 TCTGCGCGCGCTCTTACACAGCAGCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 773  
 QY 165 ThrIleTyrLysArgLysGlnLysSerGlnLeuSerAspGlyIleAlaMetLeuValAlaGly 184  
 DB 774 GAGATCCAGCAGCGACAGAG---GAGCTGACCGCAGCAGCAGCGGCTGACGAGCAG 830  
 QY 185 AsnAspArgValGlnAlaValAlaIleThrGlnMetGlnLysGlnThrIleGlnAsp 204  
 DB 831 GATAGTGCCTTTGGCGGCTTACGCGGAGATGACACGCGCGCTGCGGCGGCGGCGGCG 890  
 QY 205 AsnSerArgArgLysGlnLeuLysGlnLeuAsnGlnArgPheGlnThrLeuCysAlaValLeu 224  
 DB 891 GCGCGTCCAGCAGACCGGAGGAGCTGATCCGCGGAGCGGCTGCGCGCAGTGAAGTCAAG 950  
 QY 225 GlnGluArgLysGlnLysGlnLeuGlnAlaLeu---AlaArgGlnGlnGlu 240  
 DB 951 CGGCGTCCAGCAGCGGAGCTGCTGAGGAGCTGTGACCGCGGCTGACGAGCAGCAGCAG 1010  
 QY 241 GlnLys---LeuGlnArgValArg--- 247  
 DB 1011 GAGATGCGCAGTGGCTGCGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1070  
 QY 248 -----Gly 248  
 DB 1071 GCGCTGTCAGAGAGATGATGCTAGCCTCGACAGCAGGAGGCTGTCGACATGACCGCT 1130  
 QY 249 LeuIleArgGlnTyrGlnLysPheHisLeu---GlnGlySerSerLysLeuValGlnSerAla 267

DB 1131 TTCTGCGCGCAGGCGCTCTGCGCGCTGCGCGCGCAGGAGGAGCCCGCAGAGCTGCC 1190  
 QY 268 IleGlnSerMetGlnGluProGlnMetAlaLeuTyrLeuGlnAlaLysGlnLeuIle 287  
 DB 1191 GTGCGCAGCGAGTGGCTGCGAGGATTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAG 1250  
 QY 288 AsnLysValGlyAlaMetSerLysValGlnLeuAlaGlyArgProGlnProGlyTyrGln 307  
 DB 1251 ACCCAG---GGGAAAGATGACAGCTGTATCCAAAGAAAGCCAGCGCGCTGCGCAGCT 1307  
 QY 308 SerMetGlnGlnPheSerValSerValGlnHisValAlaGlnMetLeuArgTyrIleAsp 327  
 DB 1308 CCCAGGACCGCTATGACGTTGACCTGCGCGCGAGGAGGCGAGAGAGTGAAGGCC----- 1361  
 QY 328 PheGlnProGlyAlaAlaGlyAspGlnGlnAspAspAspMetAlaLeu 343  
 DB 1362 ---CAGTTCAGCGCGCTGCGCTGCGTGAAGCCCGCAGCTATGCTGTG 1406

RESULT 10  
 US-08-892-747-13  
 ; Sequence 13, Application US/08892747  
 ; Patent No. 6057153  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shaji T. George, Michael Ma, Martina Werner,  
 ; APPLICANT: Umberto Pace and Allan R. Goldberg  
 ; TITLE OF INVENTION: Stabilized External Guide Sequences  
 ; NUMBER OF SEQUENCES: 50  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Patrea L. Pabst  
 ; STREET: 2800 One Atlantic Center  
 ; CITY: Atlanta  
 ; STATE: Georgia  
 ; COUNTRY: USA  
 ; ZIP: 30309-3450  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/892,747  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/372,556  
 ; FILING DATE: January 13, 1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US96/00513  
 ; FILING DATE: January 19, 1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Pabst, Patrea L.  
 ; REGISTRATION NUMBER: 31,284  
 ; REFERENCE/DOCKET NUMBER: IL1109CIP2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (404) 873-8794  
 ; TELEFAX: (404) 873-8795  
 ; INFORMATION FOR SEQ ID NO: 13:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3511 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; ANTI-SENSE: NO  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: 1..3511  
 ; OTHER INFORMATION: /function= "PML-RAR DNA Sequence."

US-08-892-747-13



## Alignment Scores:

Pred. No.:	0.0753	Length:	246240
Score:	131.00	Matches:	37
Percent Similarity:	42.96%	Conservative:	21
Best Local Similarity:	27.41%	Mismatches:	53
Query Match:	6.85%	Indels:	24
DB:	2	Gaps:	6

US-09-908-988b-2 (1-366) x US-08-724-394A-20 (1-246240)

QY 22 LysGlnLeuIleCysProIleCysLeuGluMetPheSerLysProValIleLeuPro 41

Db 71769 GAGGAAGCACCCTGCTCCATGCTGACCTGATGACCAACCAAGCAATGATC---AAC 71825

QY 42 CysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnIleAsnProLeu 61

Db 71826 TGTGACACAGCTACTGCTGCTGATATACAGACTTCTTTAA-----AACCCAAAGC 71879

QY 62 TrpGlnSerArgLysSerThrValSerSerGlyArgPheArgCysProSerCys 81

Db 71880 CAAAGCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 71921

QY 82 ArgHisGluValIleuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuVal 101

Db 71922 CGGCTCCATTCATATATGAT-----AGCCTCGACCAACCAAGCAAGCTG 71966

QY 102 GluAsnIleIleAspIleTyrLysGlnGluSerSerArgProLeuHisAlaLysAlaGlu 121

Db 71967 GGAGACCTCATTTGAAGCCCTCAAGAG-----ACGGAT 71999

QY 122 GlnHisLeuMetCysGlnGluHisGluAspGluLysIleAsnIleTyrCysLeuSerCys 141

Db 72000 CAGAAGATGTCTATGTGAGGACAC---GGAGACGACTTCACCTGTCTGCGAAGACGAG 72056

QY 142 GluValProThrCysSerLeuCysLysValPheGlyAlaHisLys 156

Db 72057 GGGCAGCTCATCTGCTGGCGCTGTGAGCGGACACACAGACACAA 72101

## RESULT 12

US-08-724-394A-21

; Sequence 21, Application US/08724394A

; Patent No. 5872237

; GENERAL INFORMATION:

; APPLICANT: Feder, John N.

; APPLICANT: Krommal, Gregory S.

; APPLICANT: Laufer, Peter M.

; APPLICANT: Ruddy, David A.

; APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1

; TITLE OF INVENTION: Sequences and Antibodies Thereeto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/724,394A

; FILING DATE: 01-OCT-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitts, Renee A.

; REGISTRATION NUMBER: 35,136

; REFERENCE/DOCKET NUMBER: 017957-000100

## TELECOMMUNICATION INFORMATION:

TELEPHONE:	415-576-0200
TELEFAX:	415-576-0300
INFORMATION FOR SEQ ID NO:	21:
SEQUENCE CHARACTERISTICS:	
LENGTH:	246240 base pairs
TYPE:	nucleic acid
STRANDEDNESS:	not relevant
TOPOLOGY:	not relevant
MOLECULE TYPE:	CDNA

FEATURE: NAME/KEY: misc.feature

; LOCATION: 1..246240

; OTHER INFORMATION: /note= "HLA-H. CONTIG"

US-08-724-394A-21

QY 22 LysGlnLeuIleCysProIleCysLeuGluMetPheSerLysProValIleLeuPro 41

Db 71769 GAGGAAGCACCCTGCTCCATGCTGACCTGATGACCAACCAAGCAATGATC---AAC 71825

QY 42 CysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnIleAsnProLeu 61

Db 71826 TGTGACACAGCTACTGCTGCTGATATACAGACTTCTTTAA-----AACCCAAAGC 71879

QY 62 TrpGlnSerArgLysSerThrValSerSerGlyArgPheArgCysProSerCys 81

Db 71880 CAAAGCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 71921

QY 82 ArgHisGluValIleuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuVal 101

Db 71922 CGGCTCCATTCATATATGAT-----AGCCTCGACCAACCAAGCAAGCTG 71966

QY 102 GluAsnIleIleAspIleTyrLysGlnGluSerSerArgProLeuHisAlaLysAlaGlu 121

Db 71967 GGAGACCTCATTTGAAGCCCTCAAGAG-----ACGGAT 71999

QY 122 GlnHisLeuMetCysGlnGluHisGluAspGluLysIleAsnIleTyrCysLeuSerCys 141

Db 72000 CAGAAGATGTCTATGTGAGGACAC---GGAGACGACTTCACCTGTCTGCGAAGACGAG 72056

QY 142 GluValProThrCysSerLeuCysLysValPheGlyAlaHisLys 156

Db 72057 GGGCAGCTCATCTGCTGGCGCTGTGAGCGGACACACAGACACAA 72101

## RESULT 13

US-08-724-394A-22

; Sequence 22, Application US/08724394A

; Patent No. 5872237

; GENERAL INFORMATION:

; APPLICANT: Feder, John N.

; APPLICANT: Krommal, Gregory S.

; APPLICANT: Laufer, Peter M.

; APPLICANT: Ruddy, David A.

; APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1

; TITLE OF INVENTION: Sequences and Antibodies Thereeto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco



```

QY 123 sleuMetCysgluGluHisgluAspGluLysIleasnIleTyrCysLeuSerCysgluVal 143
Db 694 -----ATGGAGCTCTTCTGCGACAGACCGACAGAC 722
QY 143 lProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysgluValAlaProle 163
Db 723 CTGCATCTGCTACCTTTCG---ATGTTCCAGAGCACAAGATCATAGACCGTGACAGT 779
QY 163 uproThrIleTyrLysArgLysSerGluLeuSerAspGlyIleAlaMetLeuValAl 183
Db 780 GGAGGAGGCCAAGCCGAGAAAGAGACGAGCTGCA-----CT 818
QY 183 agLysAsnAspArgValGlnAlaValIleThrGlnMetGluGluValCysGln----- 200
Db 819 GCAAAAGAGCAGCTGAGCTCAAGATCATTTAGATTGAGTGAAGCTGAGAAGTGGCA 878
QY 201 -----ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnAr 216
Db 879 GAAGGAGAGAGCCGACATCAAGACCTTCACCAATGAGAAAGCCATCTGAGACAGAA 938
QY 216 gPheGluThrLeuCysAlaValIleGluGluArgLysGluGluLeuGlnAlaLeuAl 236
Db 939 CTTCGGGAGCTGCTGGGGAGCTGAGAAACAAAGAGAGAGTGAAGGCTGCGCTGGA 998
QY 236 aArgGluGlnGluLysLeuGlnArgValArgGlyLeuLeuArgGlnTyrGlyAspHi 256
Db 999 GCAAGCGGACAGATCTGCTGAGCAAGTGAAGTGAATCATG-----GATGC 1046
QY 256 sleuGluGlySerSerLysLeuValGluSerAlaIleGlnSerMetGluGlu----- 273
Db 1047 TCTGGATGAGAGAGCCAGAGTGTGCTGAGAGACAGACAGCCGCGAGAGCTGCATAG 1106
QY 274 ----ProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGluLeuIleAsnLysValGlyAl 292
Db 1107 CATCAGCAGACTGTGTGTTCTTCTGAGAA-----TTGGTGC 1145
QY 292 aMetSerLysValGlnLeuAlaGlyArgProGluProGlyTyr----- 306
Db 1146 ATTTGATGAGCAATTACTCTCTCCGCCCCCTGCCCCACCATCATCATGCTGTCGAGAG 1205
QY 307 -----GluSerMetGluGlnPhe-----SerValSerValGly 317
Db 1206 GGAAGGCGCTGGGACAGTCACTAGGCAACTTCAAGAGACGACCTGCTCATATGCATGCG 1265
QY 317 uHisValAlaGluMetLeuArg 324
Db 1266 CCACGTTGAGAAAGATGTGCAAG 1287

```

## RESULT 15

PCT-US93-05794-2

Sequence 2, Application PC/TUS9305794

GENERAL INFORMATION:

APPLICANT: The Regents of the University of California

TITLE OF INVENTION: Gene for Ataxia-telangiectasia

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Leona L. Lauder

STREET: 177 Post Street, Suite 800

CITY: San Francisco

STATE: California

COUNTRY: San Francisco

ZIP: 94108-4731

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/05794

FILING DATE: 19930618

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/903,466

```

FILING DATE: 22-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: 91-077-1 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-421-4973
TELEFAX: 415-421-1663
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1767 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1764
PCT-US93-05794-2

Alignment Scores:
Pred. No.: 7,74e-05
Score: 127.00
Percent Similarity: 40.91%
Best Local Similarity: 21.43%
Query Match: 6.64%
DB: 5
Gaps: 13

US-09-908-988b-2 (1-366) x PCT-US93-05794-2 (1-1767)
QY 46 LeuCyArgLysCysAlaAsnAspValPheGlnAla-SerAsnProLeuTyrGlnSerAr 65
Db 514 CTGCGAGCTCTGCTGAGTGGCAGCAAGAGAGAGCGGCTCAAGTCTGCTGCTGCTGCGAG 573
QY 65 g---GlySerThrThrValSerSerGlyArgPheArgCysProSerGlyArg---Hi 83
Db 574 GCCTCTTGTGAGAGTGTCAATGAGCCCAAGCTGAGAGGCGCGCTTC-CGAGACCA 632
QY 83 sGluValAlaLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuValGluAs 103
Db 633 CCAG-----CTGCTCGAGCC 647
QY 103 nIleIleAspIleTyrLysGlnLysSerSerArgProLeuHisAlaLysAlaGluGlnHi 123
Db 648 CATCGGAGCTTTGAGGCGCCGCAAGTGT-----CCGTCATGCGCAAGC----- 693
QY 123 sleuMetCysgluGluHisgluAspGluLysIleasnIleTyrCysLeuSerCysgluVal 143
Db 694 -----ATGGAGCTCTTCTGCGACAGACCGACAGAC 722
QY 143 lProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysgluValAlaProle 163
Db 723 CTGCATCTGCTACCTTTCG---ATGTTCCAGAGCACAAGATCATAGACCGTGACAGT 779
QY 163 uproThrIleTyrLysArgLysSerGluLeuSerAspGlyIleAlaMetLeuValAl 183
Db 780 GGAGGAGGCCAAGCCGAGAAAGAGACGAGCTGCA-----CT 818
QY 183 agLysAsnAspArgValGlnAlaValIleThrGlnMetGluGluValCysGln----- 200
Db 819 GCAAAAGAGCAGCTGAGCTCAAGATCATTTAGATTGAGTGAAGCTGAGAAGTGGCA 878
QY 201 -----ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnAr 216
Db 879 GAAGGAGAGAGCCGACATCAAGACCTTCACCAATGAGAAAGCCATCTGAGACAGAA 938
QY 216 gPheGluThrLeuCysAlaValIleGluGluArgLysGluGluLeuGlnAlaLeuAl 236
Db 939 CTTCGGGAGCTGCTGGGGAGCTGAGAAACAAAGAGAGAGTGAAGGCTGCGCTGGA 998
QY 236 aArgGluGlnGluLysLeuGlnArgValArgGlyLeuLeuArgGlnTyrGlyAspHi 256

```





```

: Sequence 1 Application PC/TUS9305794
: GENERAL INFORMATION:
: APPLICANT: The Regents of the University of California
: TITLE OF INVENTION: Gene for Ataxia-Telangiectasia
: TITLE OF INVENTION: Complementation Group D (ATDC)
: NUMBER OF SEQUENCES: 45
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Leona L. Lauder
: STREET: 177 Post Street, Suite 800
: CITY: San Francisco
: STATE: California
: COUNTRY: San Francisco
: ZIP: 94108-4731
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/05794
: FILING DATE: 19930618
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 07/903,466
: FILING DATE: 22-JUN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Lauder, Leona L.
: REGISTRATION NUMBER: 30,863
: REFERENCE/DOCKET NUMBER: 91-077-1 PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-421-4973
: TELEFAX: 415-421-1663
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3018 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: PCT-US93-05794-1

Alignment Scores:
Pred. No.: 0.000183 Length: 3018
Score: 127.00 Matches: 66
Percent Similarity: 40.91% Conservative: 60
Best Local Similarity: 21.43% Mismatches: 104
Query Match: 6.64% Indels: 79
DB: 5 Gaps: 13

US-09-908-988B-2 (1-366) x PCT-US93-05794-1 (1-3018)
QY 46 LeucysArglyscysAlaasnarpvalPheGlnAla-SerAsnProLeuTrpGlnSerAr 65
||| ||| :: :||| |||:|||| |||:||||
DB 638 CTTGCGACACTCTGATGAGGCAACAGAGAAGCGGTCATAGTCTGCTGTTGCGAG 6979
QY 65 g---GlySerThrThrValSerSerGlyGlyAlaArgPheArgCysProSerCysArg---H1 83
||| ||| ::| ||| |||:|||| ||| ||| ||
DB 698 GCCCTCTTCTGCGACTCATCTCAAGCCCACTGGAGGCGCGGCTTC-CGAGACCA 756
QY 83 gGUlValIleuAspArgHISGlyValTyrGlyLeuGlnArgAsnLeuValGluAs 103
|::|
DB 757 CCAG-----CTGCTCGAGCC 771
QY 103 llelleasplleTyrLysGlnGlnLuserSerArgProLeuHISAlaLysAlaGlnGlnH1 1223
||| ||| ::||| |||:|||| |||
DB 772 CATCGGGGACTTTGAGGCCCGCAAGTGT-----CCGTCGATGCGAAGAGC----- 817
QY 123 sleuMeCysGlnGlnIuHISGlnAspGluLysIleAsnIleTyrCysLeuSerCysGluVal 143
|::|
DB 818 -----ATGAGAGCTCTTTCGCACAGACCGACCGACAGAC 846
QY 143 lProthrcysSerLeuCysLysValAlpheGlyAlaHISLysAspCysGluValAlaIlePole 1635

```

```

Db      847 CTGCATCTGCTACTTTC---ATTGTCCAGACACCAAGAATCATAGCACCGTGACAGT   903
QY      163 upProthrileTyLysArgGlInySercIubSeArspgilyleAlametLeuValAl    183
        :                               ::::::::::::::::::::|
Dd      904 GGAGCAGGCCAACGGCCGAAGAGAGACGAGCTGTCA-----CT          942
QY      183 agLYsnAspaRvAgVAlnAlVaIlleThrcImmetglUvalCYsgln-----     200
        DB         GAAAAAGGACGCTCACCCTCAAAGTCAWTTAAGATTGAGGTGAAAGCTGAACAAGTCGCA   1002
QY      201 -----ThrILeLuApasnSerArgArgLnlySGlnLeuLeuanGLnar       216
        Dd      1003 GAAGGAGAAAGACCCGATCAACAGCTTCACCCACCATAGAAAGGCCATCTGTGACACAAA   1062
QY      216 gPheGIurHrLeuCYSAlAValleugLIuarqLyseLygleuleauGLnaLEuaI      236
        Db      1063 CPTCCGGACCTCGTGC CGGAGCTCGAGAAAGCAAAAGAGMAAGTAGAGGGCTGCGCTGGA   1122
QY      236 aaRGluGLInyluglULysleuGLNargVALarGValARgLYleueatrgINtyRLyghIPhi    256
        Dy      1123 GCAGCGGGAGCAGAGATGCTGTGTGACCAACAGTAAGTGATCATG-----GATGC   1170
QY      256 sLeugLIusyerSerLYSLeuVAleUGserALaleInsErMetGUgu-----      273
Dd      1171 TGTGATGACAGAACCCAAGTCTCATGAGGACAAAGACACCCGGAGAGCTGCATARG   1230
QY      274 ----ProgImEtAlAleutryLeuGLnglNaIAlysGLleuasLYsValGYAI    292
Dd      1231 CATCAGCGACTGTGTGTGTCTCAGAA-----TTTGtgc           1269
QY      292 ametserLYSVaIGLUleUAlaGLyaRProlGuProgLTYr-----             306
Dd      1270 ATTGATGAGCAATACTACTCTCTCCCCACCCCTGCCACCTATCATATGTCGTGGAGGG   1329
QY      307 -----GluseRMetGUlglnPh-----ServAlservalGI    317
Dd      1330 GGAGGCGCTGGCAGTACTACTAAGGCACCTCAAGACGACCTGTCTCATGTATGATCGG   1389
QY      317 uHIsvAlaaglumetLeuArq 324
        Dd      1390 CCACGTTGAGAAAGATGTGCAG 1411

RESULT 18
US-08-801-308-2
; Sequence 2, Application US/08801308
; Patent No. 6368790
; GENERAL INFORMATION:
; APPLICANT: Scott, Robert E.
; TITLE OF INVENTION: CDNA ENCODING P2P PROTEINS AND USE OF
; TITLE OF INVENTION: P2P CDNA-DERIVED ANTIBODIES AND ANTISENSE REAGENTS IN
; TITLE OF INVENTION: DETERMNING THE PROLIFERATIVE POTENTIAL OF NORMAL,
; NUMBER OF SEQUENCES: 4 ABNORMAL AND CANCER CELLS IN ANIMALS AND HUMANS
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weiser & Associates, P.C.
; STREET: 230 S. Fifteenth Street, Suite 500
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,308
FILING DATE: 18-FEB-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
```



Query Match: 6.41% Indels: 109  
DB: 4 Gaps: 19  
US-09-908-988b-2 (1-366) x US-08-697-610-1 (1-2240)  
QY 20 LeuGlulysGlnLeuIleCysProIleCysLeuGluMetPheSerLysProValIle 39  
DB 275 GTGGAGGACAGACAGTACAGTGTGAGAGAGTCCAGCGTGTGTGACGCCGAAGCAG-- 331  
QY 40 LeuProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnIleAsn 59  
DB 332 ACCGAGTGTGGGCGCCCTCTCTCGAGAGCTGCATGGCGCCCTCTGAGCTTCAAGT 391  
QY 60 ProLeuTrpGlnSerArgLysSerThrThrValSerSerGlyArgPheArgCysPro 79  
DB 392 CCA-----AAATGTACA 403  
QY 80 SerCysArgHisGluValAlaLeuAspArgHisGlyValTyrGly-----LeuGln 96  
DB 404 GCCTGTCAAGACAGACATCGTTAAGATAG-----CTGTTAAGATATTTGCTGCAAG 457  
QY 97 ArgAsnLeuLeuValGluAsnIleIleAspIleTyrLysGlnGluSerSerArgProleu 116  
DB 458 AGAGAATTTCTGGT-----CTTCAGATCTATTGTCCGAATGAAGACAGA----- 502  
QY 117 HisAlaLysAlaGluGln-----HisLeuMetCysGluGlu 128  
DB 503 ---GGTTGTGCAGACAGATTACGCTGGACATCTGCTGTCATTTAAATAATGATTGC 559  
QY 129 HisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValPro----- 144  
DB 560 CATTTTGAAGAA-----CTTCCATGTGTGCGTCCGCTGACGCAAGAA 601  
QY 144 ----- 144  
DB 602 AAGGCTTGAGAAAGACCTGCGACAGACAGTGAAGAAGCGGTAAATACCGGAAGCC 661  
QY 145 ThrCysSerLeuCysLys-----ValPheGlyAlaHisLysAsp--- 157  
DB 662 ACATGAGCGACACTGCAAGAGTCCGATGATCGCGTGCAGAAACAGAGACAGACCC 721  
QY 158 -----CysGluValAlaProLeuProThrIleTyrLysArgGlnLysSerGluLeu 174  
DB 722 GACTGTCCTCGCTGGTGGTCTCTGCTCCCTCAACAAGTGCAGCGCTCTCTGAG 781  
QY 175 SerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgValGlnAlaVal----- 191  
DB 782 ACCGAGGGGACAAACAGACAGATCAAGGCGCCAGCGCTCCGCCCTGACAGCAGCTC 841  
QY 192 -----IleThrGlnMetGluGluValCysGlnThrIleGluAspAsn 205  
DB 842 AACCTGCTGAAGAGTGAAGCAACTGCTCGAAAGAGAGGTTCTTGTGCAAGATGAA 901  
QY 206 SerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeuCysAla-----Val 223  
DB 902 ACTGTAGAAAAAACAAGACATACAAAGTTTCCACATCAGATATGTACTTTGAATTT 961  
QY 224 LeuGluGluArgLysGlyLysLeuLeuGlnAlaLeuAlaArgGlnGlnGluLysLeu 243  
DB 962 GAAATTTGAGAGCAAAAGAAATGCTT-----CGAATATATGATCCAAATC 1009  
QY 244 GlnArgValArgGlyLeuIleArgGlnTyrGlyAspHisLeuGluLysSerSerLysLeu 263  
DB 1010 CTTCATTTTACACGAGTGTATGACACCAAGCAAGCAAGTAAGTGAAGGCTTGCACAGAG 1069  
QY 264 ValGluSerAlaIleGlnSerMetGluGluProGlnMetAlaLeuTyrLeuGlnAla 283  
DB 1070 ATCCGCTCTCCCGGACAGACACGGGAGAGACAGACAGC-----ATGAAGACAGCGCTG 1123  
QY 284 LysGluLeuIleAsnLysValAlaMetSerLysValGlu-----LeuAlaGlyArgPro 302  
DB 1124 GAGTCCCTCCAGAACCCGGGTGACCGAGCTGAGAGCGGTGACAAAGCGCGGGGCAAGTG 1183

QY 303 GluProGlyTyrGlnSerMetGluGlnPheSerValSerValGluHisValAlaGluMet 322  
DB 1184 GCTCGGAACACAGCCCTGCTGAGTCC-----CAGCTGAGCCGCGCATGACAGATG 1234  
QY 323 LeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGluGluAspAspMetAla 342  
DB 1235 CTGAGTGTGACAGCATCCG--CTTACGCCAGCATGACCT-----GCG 1275  
QY 343 LeuAspGlyLys-----GlnGlyAsnAlaGlyLeuGluGluGluArgLeu 357  
DB 1276 CTTCAGGCTCTCGAGACAGCCGCAAGCTTACAAATGAGTGTCTGGAAGATTCGGAGCTA 1335  
QY 358 AspValProGluGluSerGlyLeuHis 366  
DB 1336 CAAGCGCGGAGACAGAGCCGTCAT 1362  
RESULT 20  
US-08-349-357-1  
; Sequence 1, Application US/08349357  
; Patent No. 6265556  
; GENERAL INFORMATION:  
; APPLICANT: Reed, John C.  
; APPLICANT: Sato, Takaaki  
; TITLE OF INVENTION: CD40 Associated Proteins  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/349,357  
; FILING DATE: 02-DEC-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 1203  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2240 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 137..1766  
; US-08-349-357-1  
Alignment Scores:  
Pred. No.: 0.000378 Length: 2240  
Score: 122.50 Matches: 86  
Percent Similarity: 44.7% Conservative: 55  
Best Local Similarity: 21.03% Mismatches: 160  
Query Match: 6.41% Indels: 109  
Gaps: 19  
US-09-908-988b-2 (1-366) x US-08-349-357-1 (1-2240)  
QY 20 LeuGlulysGlnLeuIleCysProIleCysLeuGluMetPheSerLysProValIle 39  
DB 275 GTGGAGGACAGTACAGTGTGAGAGAGTCCAGCGTGTGTGACGCCGAAGCAG-- 331



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 3, 2002, 12:12:18 ; Search time 261 Seconds  
(without alignments)  
3157.975 Million cell updates/sec

Title: US-09-908-988b-2  
Sequence: 1 MNFTVGFKPLGDHNMNDL.....EGNAGLEERLDPVEGSLH 366

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=framed\_p2n\_model -DEV=xlh  
-O=C992\_1/USPTO.spool/US09908988/runat\_26112002\_112153\_11206/app.query.fasta\_1.519  
-DB=N\_Geneseq\_101002 -ORF=fastap -SUFFIX=trng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=D10sum62 -TRANS=human4.0.cdi  
-LIST=45 -DOCMALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=20  
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09908988.ecgn.1.1.0.0runat\_26112002\_112153\_11206 -NCPu=6 -ICPU=3  
-NO\_XLPHY -NO\_MMP -FARGROUP=NEG -SCORES=0 -MATT=LONGLOC -DEV=TIMEOUT-120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FARGOP=6 -FARGEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: N\_Geneseq\_101002.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
25: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1912	100.0	1431	24	ABA99061
2	1780	93.1	1913	22	AAH68563
3	1780	93.1	1990	22	AAH78026
4	1725	90.2	1349	21	AAH72433
5	1599	83.6	2040	24	ABN85313
6	1261.5	66.0	1500	24	AAE27653
7	1162	60.8	2590	24	ABA99062
8	1147.5	58.1	1762	22	AAE25855
9	1110	58.1	1597	24	ABA99063
10	1101.5	57.6	2110	22	AAE42490
11	1098.5	57.5	1231	22	AAH80037
12	1098.5	57.5	1764	22	ABA83058
13	1091.5	57.1	1781	22	AAE25842
14	987.5	51.6	1183	22	AAH90117
15	920.5	48.1	1796	22	AAH89924
16	659.5	34.5	531	21	AACT5285
17	659.5	34.5	531	21	AACT5285
18	614	32.1	587	24	ABN78104
19	427.5	22.4	4345	23	AAE26314
20	427.5	22.4	4345	23	ABV22250
21	427.5	22.4	4345	23	ABV23763
22	427.5	22.4	4345	23	ABV24205
23	427.5	22.4	4345	23	ABV24320
24	427.5	22.4	4345	23	ABV24652
25	427.5	22.4	4345	23	ABV27751
26	427.5	22.4	4345	23	ABV28087
27	427.5	22.4	4345	23	ABV29637
28	416	21.8	446	22	AAE26303
29	411	21.5	2058	23	AAE82247
30	411	21.5	2237	23	AAE82248
31	411	21.5	6632	22	AAE82234
32	306	16.0	650	21	AACT5760
33	298	15.6	391	21	AACT5969
34	286	15.0	2272	23	AAE77751
35	275	14.4	2272	23	AAE77751
36	274	14.3	1409	22	AAE44635
37	274	14.3	1755	21	ABQ1959
38	274	14.3	1801	22	AAE16294
39	274	14.3	1801	22	AAE16294
40	259.5	13.6	1761	23	AAE5262
41	259.5	13.5	1276	24	ABQ72500
42	255.5	13.4	3262	21	AAE8383
43	234	12.2	2981	21	AACT7707
44	225	11.8	599	22	AAE16184
45	225	11.8	599	22	AAE16185

## ALIGNMENTS

RESULT 1  
ID ABA99061 standard; DNA; 1431 BP.  
AC ABA99061;  
XX  
XX  
DT 15-JUL-2002 (first entry)  
DE Murine muscle ring finger protein 1 (MURF-1) coding sequence.  
XX  
XX Muscle ring finger; MURF-1; mouse; cardiant; microtubule;  
KW intermediate filament; striated muscle; cardiac hypertrophy;  
KW heart disease; gene; ds.  
XX  
XX  
OS Mus musculus.  
XX  
XX  
FH Key Location/Qualifiers  
CDS 199..1299

```
FT      /tag="a
FT      /product="MURF-1"
XX
XX      MO200206318-A2.
XX
XX      24-JAN-2002.
XX
XX      18-JUL-2001; 2001WO-US22896.
XX
XX      18-JUL-2000; 2000US-219020P.
XX
XX      (TEXA ) UNIV TEXAS SYSTEM.
XX
XX      Olson EN, Spencer JA;
XX
XX      WPI: 2002-241506/29.
XX      P-PSDB; ABB08275.
XX
XX      Novel muscle ring finger protein useful for drug screening, and for
XX      diagnosing and treating diseases, particularly cardiomyopathies -
XX
XX      Claim 4; Page 123-125; 134pp; English.
XX
XX      The sequence encodes murine muscle ring finger protein 1 (MURF-1). The
XX      invention relates to a purified muscle ring finger (MURF) protein,
XX      selected from MURF-1, MURF-2 and MURF-3. The MURF proteins of the
XX      invention are involved in microtubule and intermediate filament
XX      stabilisation of striated muscle cells and have cardiant activity.
XX      The MURF proteins are useful for screening a candidate substance
XX      for MURF protein-binding activity, in a cell, cell-free system or
XX      in vivo, and its effect on interaction of MURF with microtubules,
XX      homodimerisation of MURF, MURF-1, MURF-2 or MURF-3 stabilisation
XX      of microtubules, interaction of MURF with intermediate filaments,
XX      e.g. desmin, vimentin and cytokeratin, and heterodimerisation of MURF.
XX      The screened compounds are useful for treating and preventing of
XX      cardiac hypertrophy and heart diseases. MURF proteins are useful as
XX      antigens to immunise animals for the production of antibodies.
XX
XX      Sequence 1431 BP; 338 A; 384 C; 473 G; 236 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.:      1.86e-178      Length:      1431
XX      Score:      1912.00      Matches:      366
XX      Percent Similarity:      100.00%      Conservative:      0
XX      Best Local Similarity:      100.00%      Mismatches:      0
XX      Query Match:      100.00%      Indels:      0
XX      DB:      24      Gaps:      0
XX
XX      US-09-908-988B-2 (1-366) x ABA99061 (1-1431)
XX
XX      1 MetAspPheThrValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20
XX      199 ATGAACCTTCACGGTGGGTTTCAAGCCGCTCTAGGGGATCGCCACACATGTGAACACTTG 258
XX
XX      21 GluLysGlnLeuIleCysProIleCysLeuGluMetPheSerLysProValIleLeu 40
XX      259 GAGAGCAGCTCATTTGCCCATCTGCTGAGATGTTTCCAAAGCCCGTGATCTTG 318
XX
XX      41 ProCysGlnHisAsnLeuLysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60
XX      319 CCCTGCCACACACACCTGTGGCGCAGTGTGCCACGACGCTCTTCCAGGCTCTTAATCTCT 378
XX
XX      61 LeuTrpGlnSerArgGlySerThrThrValSerSerGlyGlyArgPheArgCysProSer 80
XX      379 CTGTGGCAATCCCGGGGCTCCACACAGGTTCTTCAAGAGAGACGTTTCCGATGCCCATCT 438
XX
XX      81 CysArgHisGlnValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100
XX      439 TGTAGCAGCAGAGGTGTCTCGACAGAGCATGTGTCTATGGCCGTGACGGGAACCTGCTA 498
XX
XX      101 ValGluAsnIleIleAspIleTyrLysGlnGluSerSerArgProLeuHisAlaLysAla 120
XX      499 GTGGAGAACATCATTTGACATCTACACAGCAGAGAGTCTCCCGGCGACATGCAGCCAGGCT 558
```

```
QY      121 GluGlnHisLeuMetCysGlnGluHisGlnLysAspGluLysIleAsnIleTyrCysLeuSer 140
DB      559 GAACAGCAGCTCATGTGTGAGAGACACAGAGACGAAAGATCAACATCTACTGCTCCAGC 618
QY      141 CysGlnValProThrCysSerLeuLysValPheGlyAlaHisLysAspCysGlnVal 160
DB      619 TGGAGAGTGGCCACCTGCTCTCTGCAAGGTTTTCGGCGGCCACAGAGACGTGAGGTTG 678
QY      161 AlaProLeuProThrIleTyrLysArgGlnLysSerGluLeuSerAspGlyIleAlaMet 180
DB      679 GCCCTCTGGCCACCATTTACAAAGCCAGAAAGAGTGTGAGCGATGCGATCCGAGTGG 738
QY      181 LeuValAlaGlnAsnAspArgValGlnAlaValIleThrGlnMetGlnLysValCysGln 200
DB      739 CTGTGGCGGGCAATGACCCGTGTGAGAGCATGATCACCACCATGAGAGAGGTGTGCCAG 798
QY      201 ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeu 220
DB      799 ACCATTGAGACACACAGCCGACAGACAGAACCACTTTAAACCAAGAGGTTTCAGAGACCTG 858
QY      221 CysAlaValLeuGlnGluArgLysGlyGluLeuLeuGlnAlaLeuAlaArgLysGlnGlu 240
DB      859 TGGCGGCTTTTGGAGGAGCGCAGAGCGCAACGTCCTCAAGCACTGGCCGAGAGCAGAG 918
QY      241 GluLysLeuGlnArgValArgGlyLeuIleArgGlnIleArgGlyAspHisLeuGlnLysSer 260
DB      919 GAAGAAGTTGACCCGCGGCGGCTCTATCCGCGCATGACGGAGACCACTTGGAGGGCTCC 978
QY      261 SerLysLeuValGlnSerAlaIleGlnSerMetGlnGluProGlnMetAlaLeuTyrLeu 280
DB      979 TCAAGAGTGGTGGAGTCCCGCATCCAGTCCATGAGAGAGCCGACATGGCTCTCTACCTTC 1038
QY      281 GlnGlnAlaLysGlnLeuIleAsnLysValGlnAlaMetSerLysValGlnLeuAlaGly 300
DB      1039 CACAGGCAAAAGGAGCGATCAACAAAGTCCGGCGCAATGTCMAAGTGTGAGCTGGCAGGA 1098
QY      301 ArgProGluProGlyTyrGlnLysSerMetGlnGlnPheSerValSerValGlnHisValAla 320
DB      1099 CGGCGGAGCCAGGCTATGAGAGCATGAGACATTTCTGTGAGCGGTGAGACACGTGGGCC 1158
QY      321 GluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGlnLysAspAsp 340
DB      1159 GAATGTGTGGAACCATCATGACTTCCAGCGCGGCGCTGGGGATGAAGAGATGACGAC 1218
QY      341 MetAlaLeuAspGlyGlnGluGlyAsnAlaGlyLeuGlnGluLysArgLeuAspValPro 360
DB      1219 ATGGCTTTGGATGGGGAGAGAGGCAATCGGGGCTGGAGAGAGAGCGGCTGGACGTGCCA 1278
QY      361 GluGlySerGlyLeuHis 366
DB      1279 GAAAGCTCAGGCGCTGCAC 1296
XX
XX      RESULT 2
XX      AAH68563 standard; cDNA; 1913 BP.
XX
XX      AAH68563;
XX
XX      13-SEP-2001 (first entry)
XX
XX      Human protein HP03115 coding sequence.
XX
XX      Human; gene therapy; tumour; ss.
XX
XX      Homo sapiens.
XX
XX      WO200142302-A1.
XX      14-JUN-2001.
XX
XX      06-DEC-2000; 2000MO-JF08631.
```

PR 06-DEC-1999; 99JP-0346863.  
 PR 06-DEC-1999; 99JP-0346864.  
 PR 08-FEB-2000; 2000JP-0031062.  
 PR 10-FEB-2000; 2000JP-0034090.  
 PR 10-FEB-2000; 2000JP-0034091.  
 PR 14-FEB-2000; 2000JP-0035829.  
 PR 14-FEB-2000; 2000JP-0035829.  
 PR 14-MAR-2000; 2000JP-0071161.  
 PR 30-MAY-2000; 2000JP-0160851.  
 PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
 FI Kato S, Eguchi C, Saeki M.  
 XX  
 XX WPI; 2001-381646/40.  
 DR P-PSDB; MAG93278.  
 XX  
 PT Human protein originated from tumor cell line, applicable as drug,  
 PT reagent for studying intracellular protein networks and protein source  
 PT for drug screening, also encoded cDNA for gene diagnosis and gene  
 PT therapy  
 PS Claim 3; Pages 236-240; 471pp; Japanese.  
 XX  
 CC The present sequence is a human protein coding sequence. The human  
 CC protein, preferably originated from tumour cell line, is applicable as a  
 CC drug, a reagent for studying intracellular protein networks and a protein  
 CC source for screening proteins for binding low molecular weight drugs. The  
 CC human protein coding sequence is useful for gene diagnosis and gene  
 CC therapy, expression vectors and transformant cells for detection of  
 CC ligands and receptors.  
 XX  
 SQ Sequence 1913 BP; 446 A; 537 C; 612 G; 318 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2.72e-165 Length: 1913  
 Score: 1780.00 Matches: 343  
 Percent Similarity: 97.51% Conservative: 10  
 Best Local Similarity: 94.75% Mismatches: 7  
 Query Match: 93.10% Indels: 2  
 DB: Gaps: 1  
 US-09-908-988b-2 (1-366) x AAH68563 (1-1913)  
 QY 1 MetasphenrValaIGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
 Db 303 ATGAACCTCAGACAGTGGTTTCAAGCCGCTGCTAGGGAGATGCACACAGCATGCAACCTG 362  
 QY 21 GluLysGlnLeuIleCysProIleCysLeuGluMetPheSerLysProValIleLeu 40  
 Db 363 GAGAAAGCAGCTCATCTGCCCCATCTGCTGAGATGTCTCCAAACAGTGGATCCTG 422  
 QY 41 ProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
 Db 423 CCCTGCCAACACAACTGCGCCGCAAAATGTGCCAACAGAGCTTCCAGGCGCTCAATCTCT 482  
 QY 61 LeuTrpGlnSerArgLysSerThrThrValSerSerGlyIleYArgPheArgCysProSer 80  
 Db 483 CTATGGCAGTCCCGGGGGCTCCACCACTGTCTTCAAGAGCCGCTTCCGCTCCCATCG 542  
 QY 81 CysArgHisGlnValValIleuAspArgHisGlyValTyrGlyLeuGlnAlaArgAsnLeu 100  
 Db 543 TGCAGGCGTAGAGTGTCTCTGACACAGACGCGTGTCTACGCGCTGACGAAACCTGCTA 602  
 QY 101 ValGluAsnIleIleAspIleTyrLysGlnGluSerSerArgProLeuHisAlaLysAla 120  
 Db 603 GTGGAGAACATTATGACATTACAAAGCAGAGATCTCCAGAGCCGCTGCACTCCAAAGGCT 662  
 QY 121 GluGlnHisLeuMetCysGlnGluGlnHisGlnAspGluLysIleAsnIleTyrCysLeuSer 140  
 Db 663 GAGCAGCAGCATGCTGGGAGAGCATGAAGAAGAAATCATATTTTACCTCCGACG 722  
 QY 141 CysGluValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGluVal 160

Db 723 TGTGAGTGCCACCTGCTCTCTGCAAGGCTCTTGGGTGCCCAAGACATGTAAGG 782  
 QY 161 AlaProLeuProThrIleTyrLysArgGlnLysSerGlnLeuSerAspLysIleAlaMet 180  
 Db 783 GCCCAGTCCACCATTTTACAAAGCCAGAGAGATGACCTCAGCATCGCGAATG 842  
 QY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGlnLysGln 200  
 Db 843 CTGGTGGCAGGCAATGACCGCTGCAAGCAGTGCATCACAATGAGAGAGGTGTCAG 902  
 QY 201 ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArpPheGlnThrLeu 220  
 Db 903 ACTATCGAGGACAATATGCGCGAGGCAAGACAGTGTGTAAACAGAGGTTGAGACCTG 962  
 QY 221 CysAlaValLeuGluGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGln 240  
 Db 963 TGCAGAGTGTGAGAGAGCGCAAGGCTGAGCTGTCAGAGCCGCTGCGCGCAAGAG 1022  
 QY 241 GluLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyAspHisLeuGlnLysSer 260  
 Db 1023 GAGAACTGCGAGCGCGCTCCGCGCTCATCGCTCATGTAGGACACCTGGAGGCTCC 1082  
 QY 261 SerLysLeuValGluSerAlaIleGlnSerMetGluGluProGlnMetAlaLeuTyrLeu 280  
 Db 1083 TCTAAGCTGGTGGAGTCTGCCATCCAGTCCATGGAAGACCAAAATGGCGCTGATCTC 1142  
 QY 281 GlnGlnAlaLysGlnLeuIleAsnLysValGlyAlaMetSerLysValGlnLeuAlaGly 300  
 Db 1143 CAGCAGGCGCAAGGAGCTCATCATTAAGTCCGGGCGCATGTCCAAAGTGGAGTGGCAGG 1202  
 QY 301 ArgProGluProGlyTyrGlnLysMetGlnGlnPheSerValGlnHisValAla 320  
 Db 1203 CGCCCGGAGCCAGCTTATGACAGCATGAGCATTCACCTAAGGTGGAGCACTGGCC 1262  
 QY 321 GluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGlnLysAspAsp 340  
 Db 1263 GAAATGCTGCGGACCATGACTCCAGCAGCGCGCTTCCGGGAGAGAAAG-----GAG 1316  
 QY 341 MetAlaLeuAspGlyGlnGluGlnLysAlaGlyLeuGlnGlnGlnLysArgLeuAspValPro 360  
 Db 1317 GTGGCCCGCAGGAGAGAGAGGCGCGCGCGCGGAGAGAGAGCGCGGATGGGCT 1376  
 QY 361 GluGly 362  
 Db 1377 TAAGCC 1382  
 RESULT 3  
 AAH78026  
 ID AAH78026 standard; cDNA; 1990 BP.  
 XX  
 AC AAH78026;  
 XX  
 DT 26-NOV-2001 (first entry)  
 XX  
 DE Nucleotide sequence of human cardiomyopathy associated protein (CAP).  
 XX  
 KW Human; cardiomyopathy associated protein; CAP; myocardial biopsy;  
 KW left ventricle assist device; LVAD; cardiomyopathy;  
 KW ventricular tachyarrhythmia; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key location/Qualifiers  
 FT CDS 283..1359  
 FT /\*tag= a  
 FT /product= "cardiomyopathy associated protein (CAP)"  
 XX  
 PD WO200162767-A1.  
 XX 30-AUG-2001.  
 PF 26-FEB-2001; 2001WO-US05888.

```

XX 24-FEB-2000; 2000US-0184825.
PR (GENE-) GENE LOGIC INC.
XX Bednarik D, Greene J, White M;
XX WPI: 2001-570621/64.
DR P-PSDB: AAG63832.
XX
XX Nucleic acid encoding a cardiomyopathy associated protein that is
XX differentially expressed in human left ventricle assist device (LVAD)
XX myocardial biopsy samples, useful for diagnosing, preventing or
XX treating cardiomyopathy -
XX
PS Claim 2; Page 49-51; 55pp; English.
XX
XX The present sequence encodes a human cardiomyopathy associated
XX protein (CAP) that is differentially expressed in human left ventricle
XX assist device (LVAD) myocardial biopsy samples. The predicted
XX isoelectric point of CAP is approximately 12.2. The CAP polynucleotide
XX and polypeptide are useful for diagnosing, preventing or treating
XX cardiomyopathy. The treatment diminishes the occurrence of at least
XX one of the following symptoms associated with cardiomyopathy: reduced
XX ejection fraction, increased left ventricular diastolic dimension,
XX decreased ventricular wall thickness, increased atrial size, valvular
XX regurgitation, exertional intolerance or ventricular tachyarrhythmia.
XX Modulators of CAP are also useful for preventing or treating
XX cardiomyopathy.
XX
XX Sequence 1990 BP; 524 A; 535 C; 613 G; 318 T; 0 other;
S0
XX
XX Alignment Scores:
XX
XX   No. : 2, 87e-165      Length: 1990
XX   Score: 1780.00      Matches: 343
XX   Percent Similarity: 97.51%      Conservative: 10
XX   Best Local Similarity: 94.75%      Mismatches: 7
XX   Query Match: 93.10%      Indels: 2
XX                               Gaps: 1
XX
US-09-908-988b-2 (1-366) x AAH78026 (1-1990)
QY 1 MetAsnPherThrValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20
DB 283 ATGAAGCTTCACAGTGGGTTTCAAGCCGCTGATGAGGATGCACACAGCATGACAACTCG 342
QY 21 GlyLysGlnLeuLeuLysProIleCysLeuGlnMetPheSerLysProValIleLeu 40
DB 343 GAGAAAGCAGCTCATCTGCCCCATCTGCTGAGATGTCTCCAAACAGTGGTATCTTG 402
QY 41 ProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60
DB 403 CCTTGGCAACACACACTGTGTCGCCAAATGTGCCAACAGCATCTTCCAGGCTTCGAATCT 462
QY 61 LeuTrpGlnSerArgLysSerThrThrValSerSerGlyArgPheArgCysProSer 80
DB 463 CTATGGACATCCCGGGGCTCCACACAGTGTCTTCAGAGAGCCGTTCCGCTGCCCATCG 522
QY 81 CysArgHisGlnValValLeuAspArgHisGlyAlaTyrGlyLeuGlnArgAsnLeu 100
DB 523 TGCAGGCATGAGGTTCTCTGGACAGACGAGTCTTAACGGCTTCACAGCAACCTGCTA 582
QY 101 ValGlnAsnIleIleAspIleTyrLysGlnGlnSerSerArgProLeuHisAlaLysAla 120
DB 583 GTGGAGAACATTTATGCACATTTACAGCAGAGATCATCCAGGCGCTCCACTCCAGGCT 642
QY 121 GlnGlnHisLeuMetCysGlnGlnHisGlnAspGluLysIleAsnIleTyrCysLeuSer 140
DB 643 GAGCAGCAGCTCATGTGCGAGGAGCATGAGAGAGAGAAAGATATATTATCGCTGAGC 702
QY 141 CysGlnValProThrCysSerLeuCysLysValPheGlnAlaHisLysAspCysGlnVal 160
DB 703 TGTGAGTGCCTCCACTGCTCTCTCTGCAAGGTCTTCGTTGCCCAAGAGACTGTGAGTGG 762

```

```

QY 161 AlaProLeuProThrIleTyrLysArgGlnLysSerGlnLeuSerAspGlyIleAlaMet 180
DB 763 GCCCACTGCCACCATTATACAAACGCCAGAAAGCTGAGCTACAGCATGCCGATG 822
QY 181 LeuValAlaGlnAsnAspArgValGlnAlaValIleThrGlnMetGlnLysValCysGln 200
DB 823 CTGGTGCAGGCAATGACCGCGTCCAGACAGTATGATCACACAGATGAGAGAGTGTCCAG 882
QY 201 ThrIleGlnAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeu 220
DB 883 ACATATCAGGACATATACCCGAGGACAGAAAGCATGTGTAAACCAAGCTTGAAGACCTG 942
QY 221 CysAlaValLeuGlnGlnArgLysGlyLysLeuLeuGlnAlaLeuAlaArgGlnGln 240
DB 943 TCGGCACTGCTGGAGAGAGCCGAAAGGTTAGCTGCTCCAGCGCTGGCCGAGCAAGAG 1002
QY 241 GlnLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyAspHisLeuGlnLysSer 260
DB 1003 GAGAAAGCTGCAGCGCGTCCGCGGCTCATCCGTCAGTATGGGACACCTGGAGGCTCC 1062
QY 261 SerLysLeuValGlnSerAlaIleGlnSerMetGlnGlnProGlnMetAlaLeuTyrLeu 280
DB 1063 TCTAAGCTGTGGAGTCTGCATCCAGTCATGAGAAAGCCCAATGTGCGCTGTATCTC 1122
QY 281 GlnGlnAlaLysGlnLeuIleAsnLysValGlyAlaMetSerLysValGlnLeuAlaGly 300
DB 1123 CAGCAGGCCAAAGAGAGCTGATCAATTAAGTCCGGGCGCATGTCAGAGTGCAGCGAGG 1182
QY 301 ArgProGlnProGlyTyrTyrGlnSerMetGlnGlnPheSerValSerValGlnHisValAla 320
DB 1183 CGGCGGAGGACGAGCTATGAGAGCATGAGCATGAGCAATTCACCGTAAAGGTGAGACACGCGGCC 1242
QY 321 GlnMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGlnLysAspAsp 340
DB 1243 GAAATGCTGGCGACATCGACTTCAGGCGAGCGGCTTCGCGGAGAGAAAG-----GAG 1296
QY 341 MetAlaLeuAspGlyGlnGlnGlyAsnAlaGlyLeuGlnGlnLysValPro 360
DB 1297 CTGGCCCCAGAGCGAGAGAGAGGAGGCGCGCGCGGAGAGAGAGCGCGGATGGGCT 1356
QY 361 GlnGly 362
DB 1357 TAAAGC 1362

```

RESULT 4  
AAA72433  
ID AAA72433 standard; cDNA; 1349 BP.  
XX  
AC AAA72433:  
XX  
DT 19-DEC-2000 (first entry)  
XX  
DE Human nucleic acid-binding protein NuABP-52 cDNA.  
XX  
KW Human nucleic acid-binding protein; NuABP; agonist; antagonist; EST;  
KW expressed sequence tag; drug screening; recombinant expression; antibody;  
KW reproductive disorder; infertility; immunological disorder;  
KW neurological disorder; cell proliferative disorder; cancer; tumour; ss.  
XX  
OS Homo sapiens.  
XX  
PN MO200044900-A2.  
XX  
PD 03-AUG-2000.  
XX  
XX 28-JAN-2000; 2000MO-US02237.  
PF  
XX 29-JAN-1999; 99US-0117904.  
PR  
XX 29-JAN-1999; 99US-0117905.  
XX  
PA (INCY-) INCYTE PHARM INC.  
XX



PI Tang Y<sup>T</sup>, Lai P, Hillman JL, Yue H, Azimzai Y, Lu AMD, Baughn MR;  
 PI Tran B, Shih LL, Au-Young JL;  
 DR WPI: 2000-499332/44.  
 P-PSDB: AAB21048.

XX Novel nucleic acid binding proteins, used to identify agonists and  
 PT antagonists of them, for the treatment of reproductive, immunological,  
 PT neurological and cell proliferative disorders including cancer -  
 Claim 4; Page 178-179; 180pp; English.

CC Sequences AAA72382-A72436 represent cDNAs encoding novel human nucleic  
 CC acid-binding proteins (NuABPs; AAB20997-B21051). These cDNAs were  
 CC produced by extension from an appropriate EST (expressed sequence  
 CC tag) using primers designed using the EST. The invention also relates  
 CC to expression constructs, host cells and transgenic organisms comprising  
 CC a human NuABP nucleic acid, recombinant production of the human NuABPs,  
 CC and antibodies against the human NuABPs, and also to methods of  
 CC screening modulators of human NuABP activity or expression. The human  
 CC NuABPs, and their agonists and antagonists are used to treat diseases  
 CC associated with overexpression or underexpression of functional NuABPs.  
 CC Human NuABP proteins and nucleotides, and NuABP agonists and antagonists  
 CC can be used to diagnose, treat and prevent reproductive, immunological,  
 CC neurological and cell proliferative disorders. Reproductive disorders  
 CC that may be treated using compositions of the invention include  
 CC infertility, endometriosis, disruptions of the menstrual cycle and  
 CC disruptions of spermatogenesis. Immunological disorders that may be  
 CC treated include AIDS, allergies, and autoimmune disorders such as  
 CC multiple sclerosis, rheumatoid arthritis, diabetes and systemic lupus  
 CC erythematosus. Neurological disorders that may be treated include  
 CC epilepsy, neurodegenerative conditions such as Alzheimer's disease and  
 CC Parkinson's disease, prion diseases such as Creutzfeldt-Jakob disease,  
 CC and mental disorders such as schizophrenia. Cell proliferative disorders  
 CC that may be treated include a wide variety of cancers, and also  
 CC arteriosclerosis, atherosclerosis, cirrhosis and psoriasis.

XX Sequence 1349 BP; 328 A; 372 C; 428 G; 221 T; 0 other;

#### Alignment Scores:

Pred. No.: 4,36e-160 Length: 1349  
 Score: 1725.00 Matches: 337  
 Percent Similarity: 96.168 Conservative: 14  
 Best Local Similarity: 92.338 Mismatches: 8  
 Query Match: 90.22% Indels: 6  
 DB: 21 Gaps: 2

us-09-908-988b-2 (1-366) x AAA72433 (1-1349)

OY 1 MetAsnPherThrValGlyPheIysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
 DB 129 ATGAACCTTCACAGTGGGTTCAGCGCTGCTAGGGATGCACACAGCATGAGACACCTG 188  
 OY 21 GIuYsGIuLeuIleGlySerProIleGlyLeuMetPheSerIysProValIleLeu 40  
 DB 189 GAGAGACAGCTATCTGCCCCCATCTGCTGAGATGTTCTCCAAACGAGTGTATCTCTG 248  
 OY 41 ProCysGlnHisAsnLeuGlySerArgIysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
 DB 249 CCTGCGACACACACTGTGCGCCAAATGTGCCAACGACGCTTCCAGCGCTCGAATCTCT 308  
 OY 61 LeuTrpGlnSerArgIysSerThrThrValSerSerGlyIlyArgPheArgCysProSer 80  
 DB 309 CTAATGGCACTCCCGGGGCTCCACCACTGTCTTCAGAGAGCGGTTCCGCTGCCCATCG 368  
 OY 81 CysArgHisGlnValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100  
 DB 369 TCCAGGACATGAGAGTCTCTGGACACAGCGGTCTACGCGCTGCGCAAAAGTCTCTA 428  
 OY 101 ValGluAsnIleIleAspIleTyrIysGlnGluSerSerArgProLeuHisAlaTyrAla 120  
 DB 429 GTGGAGACATATATGACATTTTACAAGCAGAGAGTCAATCCAGCGCTGCAAGGCT 488

OY 121 GIuGlnHisLeuMetCysGluGluHisGluAspGluIysIleAsnIleTyrCysLeuSer 140  
 DB 489 GAGCAGCACACCTCATGTGCGAGGACGATGAAGAAGAAAGATCAATTTTCTGCTGAC 548  
 OY 141 CysGluValProThrCysSerLeuGlyValPheGlyAlaHisIysAspCysGluVal 160  
 DB 549 TGTGAGATGCCACCGCTCTCTGTCAGAGTCTCGTGGTCCACAAAGACTGTGAGAGTG 608  
 OY 161 AlaProLeuProThrIleTyrIysArgGlnIysSerGluLeuSerAspGlyIleAlaMet 180  
 DB 609 GCCCACTGCCACCATTTTACAAACGCGAAGAGTGAAGTCAAGCATGGCATCGCGATG 668  
 OY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGluGluValCysGln 200  
 DB 669 CTGGTGGCAGGCAATGACCGCGTGCAGACAGATGATACACAGATGAGAGAGTGTGCGAC 728  
 OY 201 ThrIleGluAspAsnSerArgArgIlyGlnIysGlnLeuLeuAsnGlnArgPheGluThrLeu 220  
 DB 729 ACTATCGAGACAAATAGCCGGAGGCGAAGACAGTGTAAACCAAGAGTTTGAGAGCCTG 788  
 OY 221 CysAlaValLeuGluGluIlyArgIysGlyIleLeuGlnAlaLeuAlaArgGluGlu 240  
 DB 789 TCCGCACTGCTGAGAGGACGCGAAGGCTGAGCTCTGACAGGCTGGCCGCGAGCAAGAG 848  
 OY 241 GIuYsLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyAspHisLeuGluGlySer 260  
 DB 849 GAGAGCTGCACACGGGTCCGGGCGCTCATCCGTCACATGAGGACACACCTCGAGGCCCTCC 908  
 OY 261 SerIlysLeuValGluSerAlaIleGlnSerMetGluGluProGlnMetAlaLeuTyrLeu 280  
 DB 909 TCTAAGCTGGTGGAGTCTGCTCCATCCAGTCCATGCAAGAGCCCAAAATGGCGTGTATCTC 968  
 OY 281 GlnGlnAlaIysGluLeuIleAsnIysValGlyAlaMetSerIysValGluLeuAlaGly 300  
 DB 969 CAGCAGGCGCAAGAGACTGATCAATTAAGTCCGGGCGCATGTCCAGAGTGGAGCTGGCAAGG 1028  
 OY 301 ArgProGluProGlyTyrGluSerMetGluGlnPheSerValSerValGluHisValAla 320  
 DB 1029 CGGCGGAGCCAGGCTATGAGAGCATGAGAGCAATTCACCGTAAAGGTGGAGCACTGGGCC 1088  
 OY 321 GluMetLeuArgThrIleAspPheGlnProGlyAlaAla-GlyAspGluGluAspAsp 340  
 DB 1089 GAAATGCTGCGGACCATCTCCATTCAGCCAGCGGCTTCCGGGGAGAGAAAG- ----GA 1142  
 OY 340 PmetAlaLeuAspGly-GluGluGlyAspAlaGlyLeuGluGluIlyArgLeuAspValP 360  
 DB 1143 GGTGGCCCCAGACGAGAAAGAGAGGCAACCGGGG- ----GCCGGAAGAAAAAACGGCC 1196  
 OY 360 ToGluGlySer 363  
 DB 1197 CGGATGGGGCC 1207  
 RESULT 5  
 ID ABRN85313 standard; cDNA; 2040 BP.  
 AC ABRN85313;  
 XX 30-SEP-2002 (first entry)  
 DE Human cytoskeleton-associated protein, CSAP-4, coding sequence.  
 DE Human; cytoskeleton-associated protein; CSAP; CSAP-4;  
 KW transgenic animal; antiatherosclerotic; antiproliferative; antineoplastic;  
 KW virucide; antiviral; vasotropic; cerebroprotective; neuroprotective; cytoskeletal; gene therapy; gene; ss.  
 OS Homo sapiens.  
 XX Location/Qualifiers  
 FH 291..1271  
 FT CDS  
 FT /\*tag= a

FT /product= "CSAP-4"  
XX WO200253719-A2.  
XX 11-JUL-2002.  
XX 04-JAN-2002; 2002WO-US00178.  
XX 04-JAN-2001; 2001US-260085P.  
PR 13-FEB-2001; 2001US-268554P.  
PR 14-FEB-2001; 2001US-269111P.  
PR 23-FEB-2001; 2001US-271211P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Lu DAM, Baughn MR, Yao MG, Ding L, Honchell CD, Yue H, Tang YT;  
PI Warren BA, Duggan BM, Xu Y, Walla NK, Griffin JA, Stewart EA,  
PI Gandhi AR, Khan FA, Thangavelu K, Ison CH, Azimzal Y, Hafalia AJA;  
PI Gietzen KU, Lal PG, Sanjanwala MM, Elliott VS;  
XX WPI: 2002-583611/62.  
DR P-PSDB; ABB83475.  
XX  
XX Novel isolated human cytoskeleton-associated protein for diagnosing,  
PT treating or preventing atherosclerosis, psoriasis, leukemia, epilepsy,  
PT ischemic cerebrovascular disease, cerebral neoplasms and Alzheimer's  
PT disease -  
XX  
XX Claim 5, page 154-155; 167pp; English.  
XX  
XX The present sequence is the coding sequence for a human  
CC cytoskeleton-associated protein (CSAP). The CSAP and its coding sequence  
CC are useful in the diagnosis, treatment and prevention of a cell  
CC proliferative disorder such as actinic keratosis, atherosclerosis,  
CC psoriasis, primary thrombocythemia, leukemia; a viral infection such as  
CC those caused by adenoviruses (acute respiratory disease, pneumonia),  
CC arenaviruses (lymphocytic choriomeningitis); and a neurological disorder  
CC such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral  
CC neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease or  
CC amyotrophic lateral sclerosis. The CSAP coding sequence is also useful  
CC for creating knock out or knock in humanised animals or transgenic  
CC animals to model human diseases.  
XX  
XX Sequence 2040 BP; 465 A; 584 C; 650 G; 341 T; 0 other;  
XX  
XX  
XX Alignment Scores:  
Pred. No.: 1,94e-147 Length: 2040  
Score: 1599.00 Matches: 311  
Percent Similarity: 88.67% Conservative: 10  
Best Local Similarity: 85.91% Mismatches: 7  
Query Match: 83.63% Indels: 34  
Gaps: 2  
US-09-908-988B-2 (1-366) x ABBN5313 (1-2040)  
QY 1 MetAspHeThrValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
Db ATGAACTTACAGTGGGTTTCAAGCCCTGCTAGGGGATGCACACAGCATGACAACTCG 350  
QY 21 GluLysGlnLeuLleCysProLleCysLeuGlnMetPheSerLysProValIleLeu 40  
Db GAGAAACACATCTATGCCCATCTGCTGAGATGTTCTCCAAACCAAGTGTATCTCG 410  
QY 41 ProCysGlnHisAsnLeuLysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
Db CCCATGCCAACACACATCTGCTGCGCAAAATGTCACACGACCTTCCAGGCCCTCGAATCCT 470  
QY 61 LeuTrpGlnSerArgLysSerThrThrValSerSerGlyGlyArgCysProSer 80  
Db CTATGGCAGTCCCGGGGCTCCACACACTGTGTCTTCAGAGAGCCGTTTCCGCTGCCCATCG 530  
QY 81 CysArgHisGluValIleLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100  
|||||

Db 531 TGCAGCATGAGGTTCCTCTGACAGACACGGTGTCTAGCGGCTGCAGCAACCTGCTA 590  
QY 101 ValGluAsnIleIleAspIleTyrLysGlnLysSerArgProLeuHisAlaLysAla 120  
Db GTGGAGAACATATTATCGACATTATACAGGAGGATGATCCAGCCGCTGCATCCAAAGGCT 650  
QY 121 GlnGlnHisLeuMetCysGluGlnHisGluAspGluLysIleAsnIleTyrCysLeuSer 140  
Db GAGAGACACCTCATGTGCGAGAGACATGAAGAAGAAATCAATATTACTCCGAGC 710  
QY 141 CysGluValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGluVal 160  
Db TGTGAGGTGCCCATCTGCTCTCTGCAAGGCTCTCGGCTCCACAAAGAACTGTGAGTGT 770  
QY 161 AlaProLeuProThrIleTyrLysArgGlnLysSerGluLeuSerAspLysIleAlaMet 180  
Db GCGCCACTGCGCCACCATTTTAAACAGCCAGAG----- 803  
QY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGluValCysGln 200  
Db ----- 803  
QY 201 ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeu 220  
Db 804 -----GACATATAGCCGAGGAGCAGAAAGCTTGTAAACAGAGTTTGAGAGCCTG 854  
QY 221 CysAlaValLeuGlnGluArgLysGlyLysLeuLeuGlnAlaLeuAlaArgGlnGln 240  
Db 855 TGCACATGCTCGAGAGACGCGAAGGGTGAAGCTGTCAGAGGCGCTGCGCGGAGCAAGAG 914  
QY 241 GluLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyAspHisLeuGlnLysSer 260  
Db 915 GAGAGAGTGCAGCGCGCTCCGGGCTCATCCGTAGTAGGAGGACCACTGGAGGCTCC 974  
QY 261 SerLysLeuValGlnSerAlaIleGlnSerMetGlnLysProGlnMetAlaLeuTyrLeu 280  
Db 975 TCTTAAGCTGGTGGAGTGTCCATCCATGCATGGAAGGCCAATAATGGCGCTGTATCTC 1034  
QY 281 GlnGlnAlaLysGlnLeuIleAsnLysValGlyAlaMetSerLysValGlnLeuAlaGly 300  
Db 1035 CAGAGGCGCAAGGAGCTGATCAATATAGTCCGGGCGCATGTCAGAGGTGAGCTGGGAGG 1094  
QY 301 ArgProGluProGlyTyrGluSerMetGlnLysPheSerValSerValGlnHisValAla 320  
Db 1095 CGGCGGAGGCGAGGCTATGAGAGCATGAGACATTCACCGTAAAGGTGAGACACGTGGCC 1154  
QY 321 GlnMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGluLysAspAsp 340  
Db 1155 GAATGCTGCGGACCATCTGACCTTCCAGCCAGCGCTTCCGGGAGGAGAAAG-----GAG 1208  
QY 341 MetValLeuAspGlyGlnGluGlyAsnAlaGlyLeuGlnGluArgLeuAspValPro 360  
Db 1209 GTGGCCCCAGACGAGAGAGAGGAGGCGCGGGCCGGAGGAAGAGCGCGGATGGGCTT 1268  
QY 361 GluGly 362  
Db 1269 TAAGGC 1274  
RESULT 6  
ID AAF27653 standard; DNA; 1500 BP.  
XX AAF27653:  
AC AAF27653:  
XX  
XX 30-MAR-2001 (first entry)  
DE DNA encoding human transcriptional regulator protein #22.  
XX Human; transcriptional regulator protein; TXREG; ds.  
XX Homo sapiens.  
XX  
XX WO200078954-A2.  
PN

```

XX 28-DEC-2000.
PD
XX
XX 15-JUN-2000; 2000WO-US16766.
PF
XX 18-JUN-1999; 9905-0140109.
PR
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX
XX Lal P, Yue H, Tang YT, Baughn MR, Azimzai Y, Tran B;
XX WPI: 2001-041425/05.
XX
XX Isolated polypeptide with a human transcriptional regulator protein
XX sequence is useful for the diagnosis, prevention and treatment of
XX disorders associated with the immune, reproductive and cardiovascular
XX systems.
XX
XX Claim 5; Page 135-136; 142pp; English.
XX
XX The present invention relates to human transcriptional regulator
XX protein (TXREG) sequences. The antagonist and an agonist of the proteins
XX of the invention are used to treat disorders associated with decreased
XX or increased expression or activity of TXREG.
XX
XX Sequence 1500 BP; 401 A; 378 C; 439 G; 282 T; 0 other;
SQ
XX
XX Alignment Scores:
XX Pred. No.: 2,11e-114 Length: 1500
XX Score: 1261.50 Matches: 251
XX Percent Similarity: 89.04% Conservative: 17
XX Best Local Similarity: 83.39% Mismatches: 27
XX Query Match: 65.98% Indels: 6
XX DB: 22 Gaps: 4
XX
XX US-09-908-988b-2 (1-366) x AAF27653 (1-1500)
XX
XX 1 MetasphehrhvalgllyphelysProleuLeuGlyAspAlaHisAsnMetAspAsnLeu 20
XX
XX 291 ATGAACCTTCACAGTGGGTTCACAGCGCGCTAGAGGATGACACACAGCATGACACACCTG 350
XX
XX 21 GluysglnleuilecysProIleecysleuGluMetPheSerlyProValAlileu 40
XX
XX 351 GAGAGGACGCTCATCTGCCCCATCTGCCGAGATGTTCCAAACACAGTGGATCTG 410
XX
XX 41 ProCysGlnHisAsnLeuCysArglyscysAlaAsnAspValPheGlnLasSerAsnPro 60
XX
XX 411 CCCGCCACACACACACCTTGCCTGCGCAATGTGCCAAGACAGCTTCCAGGCGCTGGAATCCT 470
XX
XX 61 LeuTrpGlnSerArgglySerThrThrValSerSerGlyGlyArgPheArgCysProSer 80
XX
XX 471 CTATGGCAGTCCCGGGGCTCCACCACTGCTCTCAGAGGCGGCTTCCGCTGCCATCG 530
XX
XX 81 CysArgHisGlnValAlaLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100
XX
XX 531 TGCAGGCGCTGAGGTGTCTCTGACACAGACAGGCTGTCTACGGCTCGAGGAAACCTGCTA 590
XX
XX 101 ValGlnAsnIleIleAspIleTyrLysGlnGlnSerSerArgProleuHisAlaLysAla 120
XX
XX 591 GTGAGAAACATTATGACATTACAAAGCAGAGTATCCAGGCCGCTGCACTCCAGAGCT 650
XX
XX 121 GluGlnHisLeuMetCysGlnGluHisGlnAspGlnLysIleAsnIleTyrCysLeuSer 140
XX
XX 651 GAGCAGCAGCTCATGTGTGCGAGAGCATGAAGAAGAAATCAATTTACAGCCGAGAG 710
XX
XX 141 CysGlnValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGlnVal 160
XX
XX 711 TGTAAGGTGCCACCTGCTCTCTGTGCAAGGTCTTGGGCGCCCAAGAGCAGTGTGAGGTG 770
XX
XX 161 AlaProleuProThrIleTyrLysArgGlnLysSerGlnLeuSerAspGlyIleAlaMet 180
XX
XX 771 GCCCAGTGGCCACCATTTACAAAGCCAGAAAGAGTGAAGCTCAGCATGCGCATCGGATG 830
XX

```

```

QY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGlnGluValCysGln 200
DB CTGTGGCAGGCAATGACCGCGCTGCAAGCATGATCACAATGAGAGAGTGTCCAG 890
QY 201 ThrIleGlnAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeu 220
DB ACTATCGAGGACAAATGACCGCGAGCAGAAAGCAGTGTGTAAACAGAGGTTTGAAGCCTG 950
QY 221 CysAlaValAlleuGlnGluArgLysGlyGlnLeuLeuGlnAlaLeuAlaArgGlnGln-G1 240
DB 951 TGCAGAGTCTGAGAGGACGCGACGAGTGTGCTGCAAGCGCTGCGCGCGGCAAGCA 1010
QY 240 uGlnLysLeuGlnArg---ValArgGlyLeuIleArg---GlnTyrGlyAspHisLeuG1 258
DB 1011 GGACAGCTTCAACCGCATCCGAGCGGACTCATTCGCGTCAATATGCGGAGCAACTT 1070
QY 258 uGlySerSerLysLeuValGlnSerAlaIle-----GlnSerMetGlnLysProGlnMe 276
DB 1071 GGGAAGGCCCTCCCTGCTTAAGAGCGTGTGAGAGGCTGACACTGTCACCAATGCCAGTTT 1130
QY 276 tAlaLeuTyrleuGlnGlnAlaLysGlnLeuIleAsnLysValGlyAlaMetSerLysVa 296
DB 1131 CCACATTTGGATATACACAGAGC---CCAACAGAAATTTGTGTGCGGACACTTAGTAAAT 1187
QY 296 1 296
DB 1188 G 1188
XX
XX RESULT 7
XX ABA99062
XX ID ABA99062 standard; DNA; 2590 BP.
XX
XX ABA99062;
XX
XX 15-JUL-2002 (first entry)
XX
XX Murine muscle ring finger protein 2 (MURF-2) coding sequence.
XX
XX Muscle ring finger; MURF-2; mouse; cardiac; microtubule;
XX intermediate filament; striated muscle; cardiac hypertrophy;
XX heart disease; gene; ds.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX CDS 80..1717
XX FT /*tag= a
XX FT /product= "MURF-2"
XX
XX MO200206318-A2.
XX
XX 24-JAN-2002.
XX
XX 18-JUL-2001; 2001WO-US22896.
XX
XX 18-JUL-2000; 2000US-219020P.
XX
XX (TEXA ) UNIV TEXAS SYSTEM.
XX
XX Olson EN, Spencer JA;
XX
XX WPI: 2002-241506/29.
XX
XX P-PSDB; ABB08276.
XX
XX Novel muscle ring finger protein useful for drug screening, and for
XX diagnosing and treating diseases, particularly cardiomyopathies
XX
XX Claim 4; Page 126-129; 134pp; English.
XX
XX The sequence encodes murine muscle ring finger protein 2 (MURF-2). The
XX invention relates to a purified muscle ring finger (MURF) protein,
XX selected from MURF-1, MURF-2 and MURF-3. The MURF proteins of the
XX invention are involved in microtubule and intermediate filament

```

CC stabilisation of striated muscle cells and have cardiant activity.  
 CC The MURF proteins are useful for screening a candidate substance  
 CC for MURF protein-binding activity, in a cell, cell-free system or  
 CC in vivo, and its effect on interaction of MURF with microtubules,  
 CC homodimerisation of MURF, MURF-1, MURF-2 or MURF-3 stabilisation  
 CC of microtubules, interaction of MURF with intermediate filaments,  
 CC e.g. desmin, vimentin and cytokeratin, and heterodimerisation of MURF.  
 CC The screened compounds are useful for treating and preventing of  
 CC cardiac hypertrophy and heart diseases. MURF proteins are useful as  
 CC antigens to immunise animals for the production of antibodies.

XX Sequence 2590 BP; 703 A; 627 C; 625 G; 635 T; 0 other;

#### Alignment Scores:

Pred. No.:	2.78e-104	Length:	2590
Score:	1162.00	Matches:	226
Percent Similarity:	76.99%	Conservative:	55
Best Local Similarity:	61.92%	Mismatches:	76
Query Match:	60.77%	Indels:	8
	24	Gaps:	4

US-09-908-988b-2 (1-366) x ABA99062 (1-2590)

QY 1 MetasphethrValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
 DB 80 ATGAGCACTTCTCTGAATTACAAAGCTTTCTCCAAAGACGACGACCATGATTAACCTTG 139  
 QY 21 GtutysGlnLeuIleCysProIleCysLeuGluMetPheSerIlyProValIleLeu 40  
 DB 140 GAAAGCAACTGATGTGTCCTGCTCCAGAGATGTTCCAGAACTGCTGCTATTCCTC 199  
 QY 41 ProCysGlnHisAsnLeuCysArgIlyCysAlaAsnAspAlaPheGlnAlaSerAspPro 60  
 DB 200 CTTTCCACGACACACTGTCGAGAAATGTCCTGACATCTTCCAGGCTTCAACCCG 259  
 QY 61 LeuTrpGlnSerArgIlySerThrIleValSerSerIlyLysArgPheArgCysProSer 80  
 DB 260 TACTTACCCACAAAGAGAGACGACACACGTCGACAGGGGCGCTTCCGCTGTCCTCC 319  
 QY 81 CysArgHisGlnValValLeuAspArgHisGlyValIlyTrpIlyLeuGlnArgAsnLeu 100  
 DB 320 TGCACACATGAGTGGTGTAGACACAGATGGGTCTATGACATGACAGAGAACCTGCTC 379  
 QY 101 ValGluAsnIleIleAspIleTrpIlyGlnIleSerSerIlyProLysHisAlaLysAla 120  
 DB 380 GTGGAACATTAATGATCTACACAGCAAGAAATCCACAGGCCA---GAAAAAATG 436  
 QY 121 GlnGlnHisLeuMetCysGlnIleHisGluAspGluLysIleAsnIleTrpCysLeuSer 140  
 DB 437 GACCAAGCCC---ATGTGTGAAGAGCATGAAGAGACGATCAACATCTATTGCTGTAAC 493  
 QY 141 CysGlnValProThrCysSerLeuCysLysValPheGlnIleHisLysAspCysGlnVal 160  
 DB 494 TGTGAAGTCCCAACCTGTTCTGCAAGGTTTGTGCCCCCAATTAAGGACTGTGCGAGTG 553  
 QY 161 AlaProLeuProThrIleTrpIlyArgGlnIlySerSerIlyLeuSerAspGlyIleAlaMet 180  
 DB 554 GCTCCCTGACATCATCTGTTCCAGAGCAGACATCAAGAGCTACATGATGATGCTGTA 613  
 QY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGluGlnValCysGln 200  
 DB 614 CTGTGGGAGAACGATAGAGTCCAGGCTGATGACCCAGCTGAGAGACACCTGTAAA 673  
 QY 201 ThrIleGlnAspAsnSerArgIlyGlnIlyGlnLeuLeuAsnGlnIleArgPheGlnThrLeu 220  
 DB 674 ACTATTGAGAGAGTCTGTCGAAAGCAGAAACAGACCTGTGTGAGAAATTTGACCTCACTA 733  
 QY 221 CysAlaValIleGlnGluIlyArgIlyGlnLeuLeuGlnIleAlaValArgGlnGlnIle 240  
 DB 734 TACGCGATCTGAGAGAGAGAAAGACTGAATGACCCAAAGCCATCACTGACAAACAGAG 793  
 QY 241 GtutysLeuGlnArgValArgGlyLeuIleArgGlnIlyTrpIlyAspHisLeuGlnIlySer 260

DB 794 GAGAACTGGAACTATGCCAAGCTTATGACGAAGTATTCGATCCGATGAGAACGTA 853  
 QY 261 SerIlyLeuValGlyIleSerAlaIleGlnSerMetGluIlyProGlnMetAlaLeuTrpLeu 280  
 DB 854 TCCAACTGGTGGAGTCAGGAATCCAGTTCATGATGATGAGCCGAAATGGCAGTATTTCTG 913  
 QY 281 GlnGlnAlaIleGlnLeuIleAsnIlyValIlyAlaMetSerIlyValIleLeuAlaGly 300  
 DB 914 CAGATGCCAAGACCTGTTGCAAAAGATCGTGAAGCATCAAAAGCGCTTACAGTGGAG 973  
 QY 301 ArgProGluProGlyTrpIlyGlnSerMetGluIlyPheSerValSerValGlnIleValAla 320  
 DB 974 AAACATGACCAAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1033  
 QY 321 GluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGluGlnAspAsp 340  
 DB 1034 AAATTAATCCGTGAATATGACTTT-----TCTAGAGAGAGGAGGAGGAGCAAGAT 1084  
 QY 341 MetAlaLeuAspGlyGluGlnIlyAsnAlaGlyLeuGlnIlyGlnIlyArgLeuAspValPro 360  
 DB 1085 GCAGGAAATATGATGAG 1135  
 QY 361 GtutysSerGlyLeu 365  
 DB 1136 GAGGCGAGAAATGTT 1150  
 RESULT 8  
 AAS25855  
 ID AAS25855 standard; cDNA, 1762 BP.  
 AC AAS25855;  
 XX 07-NOV-2001 (first entry)  
 DE Human cDNA encoding a novel secreted protein, Seq ID 34.  
 XX  
 KW Human; immunosuppressive; antiarthritic; ss; antirheumatic;  
 KW cytosolic; cardiant; vasotropic; cerebroprotective; nootropic;  
 KW neuroprotective; antibacterial; virucide; fungicide; opthalmaological;  
 KW vulnary; secreted protein; rheumatoid arthritis;  
 KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;  
 KW cerebrovascular disorder; cerebral ischemia; anglogenesis;  
 KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;  
 KW corneal infection; wound healing; epithelial cell proliferation;  
 KW skin aging; food additive; preservative; antiproliferative.  
 OS Homo sapiens.  
 XX  
 PN WO200155322-A2.  
 XX  
 PD 02-AUG-2001.  
 PF 17-JAN-2001; 2001MO-US01341.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.



disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. The present sequence encodes a novel secreted protein of the invention.

**Alignment Scores:**

Pred. No.:	4,38e+103	Length:	176
Score:	1147.50	Matches:	25
Percent Similarity:	75.34%	Conservative:	50
Best Local Similarity:	61.64%	Mismatches:	72
Query Match:	60.02%	Indels:	18
DB:	22	Gaps:	4

US-09-908-988B-2 (1-366) x AAS25855 (1-1762)

OY		1	MetsunherthaValGlyPheLysProLeuLeuGlInsprLaHisAsnMetAspSerLeu	20
Dp		157	ATGAGGCCAATCTGTGAATTTCMAAATCTTTTCCAAAGACGACGACCATGGATACTTA	216
OY		21	GILySLInLeuIleCysProIleCysLeuGIuwerPheSerIysProVal-IleLe	40
Dp		217	GAGAACGACCTCATCTGCCCACTCGCTTGAGAATGTTCTCACGAACCTGTGTAGATTCT	276
OY		40	wProCysGLInHisAsnLeuCySarArgLysCysAlaAsnAspValPheGlnAlaSerAsnPr	60
Dp		277	CCCTTGACACACAACCTGTGTAGGAATGTGCAGAGTAATTTTCCAGGCCCTCTAACCC	336
OY		60	oleuTPGIInSerArgIlySerThrThraIleSerSergLIyARPheargCysProse	80
Dp		337	GTAATTGCCACAAGAGGAGTAGACACCAATGGCATCAGGGGGCCGATCCGCTGCCATC	396
OY		80	rCYsarHISgluValValLeuAsparHISglValTyrgLYeugInArgrasNleue	100
Dp		397	CTGTAGCAATGAAGTGGTTTTGTGATGACATGGGGTATATATGACTTCAGAGAACCTGCT	456
OY		100	uVALgluAsnIleIleAspIleTyrrLySGInIuseSerArgProLeuHisAlaLysAl	120
Dp		457	GGTGGAAAATATATTGATTCATCTACACAGCGAGAGTCCACCAAGCCA----GAAAAAGAA	510
OY		120	aGIuGINHisLeuMetCysgluGIuInHISgluaSPrgLIyLSIleasnIleTyrcYSLeusc	140
Dp		511	ATCCGACACACCCCATTGGCGAGGACATGAAAGAGGAGCGCATCAACATCACTGCTCGAA	570
OY		140	rCYsgluValProthrCYserIeucYslySvalPheglYalAHISlysAspcysgluVa	160
Dp		571	CTCGGAAGTACCCACCTGCTCTGTGCAAGGTGTGGTGGCCACACAAGAGACTGCCAGGT	630
OY		160	LAIAProLeuProthrIleTyrrLysArgrGlnLysSergIuLeuSerAspLIyleAlame	180
Dp		631	GGGTCCCCCTACATCATGTGTCCAGAGACAGAAAGCTGAGCTTAGTAATGGCATGCCAT	690
OY		180	tLeuValAlagLYAsnAspArgValGlnAlaValIleThrGImetCylgluValcysgl	200
Dp		691	CCTCGCTGGGACGACAAAGATCGAGTCCAGGAGATGATCAGCCACAGTGAAGACACGTCAA	750
OY		200	nPIrrIleGIuAsParSereArgArgIInLysGlnLeuLeuAsnGlnArgrPhegluThrie	220
Dp		751	AACATATCGAAGAAATGTTCGCAAAACAGAAACCAAGCTTTGTGAAGATTTTGATTAACT	810
OY		220	uCYsaIlaValleuGIuIwArgLysGLyuLeuLeuGlnAlaLeuAlaArgLuInGI	240
Dp		811	GTAATGCCATTTTGGAGAGAGGAAGAAATGAATAAGACCACCAAGTCATTATCCGGAACCAAGA	870
OY		240	uGIuLYSLeuGlnArgrValArgGLyLeuIleArgIntYrgLIyASPIHisLeuGIlySe	260
Dp		871	GGAGAAACAGGAACTGTCCTGCTCTGTATCAAAAGATTTTCATTCATTGGAGAACT	930
OY		260	rSerLYSLeuValGIuSerAlaIleGlnSerMetGIuIuProGImetalAuTYrIe	280

Db	931	CTCAAGTGTGGTTGACTCAGGAATTCAGTTTATGAGTAGACCAGAAATGGCAGGTTCCT	990
QY	280	UGlncGlnAlaIaIySGlUleuIleasnIySaIaGlyAlaMetSerIySaIaGlUleuAlaGl	300
Db	991	GGAGATGGCCAAACCCCTCTCTAAAAAAATCTCAGAACATCAAGGCATTTCAGATGA	1056
QY	300	YARpProGluPProGlyTyrGluSerMetGluGlnPheSerValSerValGluHsValAl	320
Db	1051	GAATAATAGACACATGGCTATAGAGACATGAGACCACTTCACAGTCAACCTCAATGACAGA	1110
QY	320	aGluMetLeuAryThrIleaspPheGlnProGlyAlaAlaGlyAspGluGluAspAspAs	340
Db	1111	AAAGATTAATACGTGAATAATTGACCTTTAC-----AGAGAAAGTGAAGA	1152
QY	340	pMetAlaIeuAspGlyGluGluGlyAsnAlaGlyLeuGluGluGluArgLeuAspValPr	360
Db	1153	T-----GAAGAGAGAAAGAACGCGGAGAGAGAAAGA-----	1186
QY	360	oGluGlySerGly	364
Db	1187	-GAAGGAGAGAGA	1198

RESULT 9  
AD2000063

ID ABA99063 standard; DNA; 1597 BP.

AC ABA99063;

DT 15-JUL-2002 (first entry)

Murine muscle ring finger protein 3 (MURF-3) coding sequence.

KW Musclevring finger; MURF-3; mouse; cardiact; microtubule;  
KW intermediate filament; striated muscle; cardiac hypertrophy;  
KW heart disease; gene; ds.

OS Mus musculus.

FH	Key	Location/Qualifiers
22		22

FT / \*tag= a

**XX**

XX 34 MAY 2003

XX  
DE 19 THH 0001 0001110 0000000

[illegible]XX  
PA (TEVA) FINTV MEVAC SYSTEM

PT	0150Z	FM	CASUALTY	TR.
XX				
XX				

XX  
XX  
DB EBT: 2003-241E0E/200

DR P-PSDB; ABB08277.  
XX

**Novel muscle ring finger protein useful for drug screening, and for diagnosing and treating diseases particularly associated with**

XX  
PS  
Claim 4: Page 131-133; 134nn; English

the sealant should be mixed with the filler according to the instructions (Figure 3). The

CC The sequence encodes the muscle ring finger protein-3 (MURF-3). The  
CC invention relates to a purified muscle ring finger (MURF) protein,  
CC selected from MURF-1, MURF-2 and MURF-3. The MURF proteins of the  
CC invention are involved in microtubule and intermediate filament  
CC stabilisation of striated muscle cells and have cardiac activity.  
CC The MURF proteins are useful for screening a candidate substance  
CC for MURF protein-binding activity. In a cell, cell-free system or  
CC *in vivo*, and its effect on interaction of MURF with microtubules,  
CC homodimerisation of MURF, MURF-1/MURF-2 or MURF-3 stabilisation







CC encodes is useful for the diagnosis and/or prognosis of one  
 CC or more types of cancer. The polynucleotide and polypeptide can be  
 CC used as nutritional sources or supplements and in the screening of  
 CC chemical compounds as potential drugs.

50 Sequence 1231 BP; 332 A; 305 C; 361 G; 233 T; 0 other;

## Alignment Scores:

Alignment No.:	1-78e-98	Length:	1231
Score:	1098.50	Matches:	210
Percent Similarity:	78.00%	Conservative:	63
Best Local Similarity:	60.00%	Mismatches:	66
Query Match:	57.45%	Indels:	11
DB:	22	Gaps:	4

US-09-908-988B-2 (1-366) x AAH90037 (1-1231)

```

QY 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluLysGlnLeuIleCysProIleCys 29
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 167 CTGATCCAGAGTGGAGATCCCATGGAACTTGGAGAACAGCAGCTGATCCCTATCGC 226

QY 30 LeuGluMetPheSerIleProValAlaIleLeuProCysGlnHisAsnLeuCysArgLys 49
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 227 CTGGAGATGTTTACCAGAGCAGTGTCTTCTGCGCCAGCAGACACCTGTCGGAG 286

QY 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTyrPglSerArgLysThrThr 69
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 287 TGTGCCATGACATCTCTCCAGGCTGCMAATCCCTACTGAGCAGCCGGGCGACACTCAGTG 346

QY 70 ValSerSerGlyLysArgPheArgCysProSerCysArgHisGluValLeuAspArg 89
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 347 TCCATGTCTGGAGCGCTTCCGCTGCCACCTCCGACAGAGGTGATCATGATGATCGT 406

QY 90 HisGlyValTyrGlyLeuGlnArgAsnLeuValGluAsnIleLeuAspIleTyrLys 109
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 407 CACGAGATGAGGCTGCTGAGAGAACCTGCTGCTGAGAAACATCTCATCTTCAAAA 466

QY 110 GlnGlu---SerSerArgProLeuHisAlaLysAlaGluGlnHisLeuMetCysGlu 128
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 467 CAGGAGTGTCTCCAGTCCGCGCTG-----CAGAAAGGCGAGTCACCCCATGTGCAGAGAG 520

QY 129 HisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValAlaProLeuProThrIleTyrLys 148
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 521 CACGAGATGAGAAATGAAATCACTACTGCTCTCAGCTGAGAGTGCCTCCTCATG 580

QY 149 CysLysValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIleTyrLys 168
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 581 TGCAAGGTGTTGGATGATCACAAGGCTCGAGGTGGCCCATTCAGAGATGCTTCCAG 640

QY 169 ArgGluLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGluAsnAspArgVal 188
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 641 GCACAAAAGATGAACTGAATTAATCTGATCTCATGCTGCTGGCGGAGATGACCTGTG 700

QY 189 GlnAlaValIleThrGlnMetGlnGluValCysGlnThrIleGluAspAsnSerArgVal 208
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 701 CAGACCATCATCTACCTCAGCTGAGGATTCCTCCGTGATGCCAGCAAGAGAACTTCAACCCAG 760

QY 209 GlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeuCysAlaValLeuGluGluValGlyLys 228
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 761 GTAAAGGAGAGAGCTGAGGACAGATTTCACACGTTGTATCCATCTGATGAGAAAGAA 820

QY 229 GlyLysLeuLeuGlnAlaLeuAlaArgGlnGluGluGluLysLeuGlnArgValArgLys 248
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 821 AGTAGAGTGTCTGACAGCGGATCAGCAGAGAGAGAGAAAGAAAGCTTACCTCATCGAGGCC 880

QY 249 LeuIleLeuArgGlnTyrGlyAspHisLeuGlnGluSerSerLysLeuValGluSerAlaIle 268
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 881 CTCATCCAGCAGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 940

QY 269 GlnSerMetGlnGluProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGlnLeuIleAsn 288
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 941 CAGTCCCTGGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1000
  
```

```

QY 289 LysValGlyAlaMetSerLysValGluLeuAlaGlyArgProGluProGlyTyrGluSer 308
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 1001 AGCATTTGGGAAGCTTCCAAAGGCTGCCAGCTGGGAGAACAGAGAGGCTTTGAGAAC 1060

QY 309 MetGluGlnPheSerValSerValGluHisValAlaGluMetLeuArgThrIleAspPhe 328
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 1061 ATGCACTTCTTACTTGTGATTTAGAGCAGCATAGAGAGAGAGAGAGAGAGAGAGAGAG 1120

QY 329 GlnProGlyAlaAlaGlyAspGluGluAspAspMetAla-----LeuAsp 344
    |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...|||
Db 345 GlyGluGluGlyAsnAlaGlyLeuGluGlu 354

Db 1169 GAGGAGAGTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1198
  
```

## RESULT 12

ABA83058 standard; DNA; 1764 BP.

ABA83058;

05-FEB-2002 (first entry)

Human transcription factor TRFX-85 coding sequence.

Human; transcription factor; TRFX; cell proliferative disease;

autoimmune disease; inflammation; neurological disease;

developmental disorder; cancer; AIDS; infection; cytostatic; anti-HIV;

neuroprotective; antiinflammatory; gene therapy; ds.

Homo sapiens.

WO2001/72777-A2.

04-OCT-2001.

13-MAR-2001; 2001WO-US08117.

13-MAR-2000; 2000US-0188986.

(INCY-) INCYTE GENOMICS INC.

Hillman JL, Baughn MR, Yue H, Lal P, Lu DAW, Patterson C;

Azimzai Y, Bandman O, Tang YT, Mathur P, Shah P, Au-Young J;

Reddy R;

WPI; 2001-570896/64.

P-PSDB; ABB50234.

Novel transcription factor polypeptides, used to treat diseases

associated with altered activity and expression of TRFX, and to screen

for agents capable of modulating its activity -

Claim 11; Page 311; 327pp; English.

The present sequence is the coding sequence for a human transcription

factor. The transcription factor and its coding sequence are useful in

the diagnosis, treatment and prevention of diseases associated with

altered expression of the transcription factor e.g. cell proliferative,

autoimmune/inflammatory, neurological and developmental disorders. A

number of specific disorders/diseases are given in the specification,

including: arteriosclerosis, cirrhosis, hepatitis, cancers, AIDS,

allergies, anaemia, asthma, autoimmune thyroiditis, bronchitis, atopic

dermatitis, diabetes mellitus, emphysema, Goodpasture's syndrome, gout,

grave's disease, multiple sclerosis, osteoarthritis, pancreatitis,

psoriasis, rheumatoid arthritis, systemic lupus erythematosus, ulcerative

colitis, uveitis, Alzheimer's disease, Huntington's disease, Parkinson's

disease, stroke, and viral, bacterial, fungal and protozoal infections.

Sequence 1764 BP; 473 A; 399 C; 516 G; 376 T; 0 other;

Alignment Scores:

Pred. No.:	2,92e-98	Length:	1764
Score:	1098.50	Matches:	210
Percent Similarity:	78.00%	Conservative:	63
Best local Similarity:	60.00%	Mismatches:	66
Query Match:	57.45%	Indels:	11
DB:	22	Gaps:	4

US-09-908-988B-2 (1-366) x ABA83058 (1-1764)

[illegible]

QY	329	GlnProGlyAlaAlaGlyAspGluGluAspAspMetAla-----LeuAsp	344
Db	1095	-----GGGACGAGATGAGGAGAGGAGAGCAATTCATTGAGAGAGAAAGATCAGGAA	1142
QY	345	GlyGluGluGlyAsnAlaGlyLeuGluGlu	354
Db	1143	GAGGAGAGAGTCCACAGAGGAGGAAGAA	1172
RESULT 13			
AS25842	standard; CDNA; 1781 BP.		
AS25842			
AC	AS25842;		
AD	07-NOV-2001	(first entry)	
DE	Human CDNA encoding a novel secreted protein, Seq ID 21.		
XX	Human; immunosuppressive; antiarthritic; ss; antirheumatic;		
KW	cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;		
KW	neuroprotective; antibacterial; virucide; fungicide; ophthalmologic;		
KW	vulnerable; secreted protein; rheumatoid arthritis;		
KW	hyperproliferative disorder; cardiovascular disorder; cardiac arrest;		
KW	cerebrovascular disorder; cerebral ischaemia; angiogenesis;		
KW	neuro system disorder; Alzheimer's disease; infection; ocular disorder		
KW	corneal infection; wound healing; epithelial cell proliferation;		
KW	skin ageing; food additive; preservative; antiproliferative.		
OS	Homo sapiens.		
XX	WO200155322-A2.		
XX	02-AUG-2001.		
PD	17-JAN-2001; 2001WO-US01341.		
XX	31-JAN-2000; 2000US-0179065.		
PR	04-FEB-2000; 2000US-0180628.		
PR	24-FEB-2000; 2000US-0184664.		
PR	02-MAR-2000; 2000US-0186350.		
PR	16-MAR-2000; 2000US-0189874.		
PR	17-MAR-2000; 2000US-0190076.		
PR	18-APR-2000; 2000US-0198123.		
PR	19-MAY-2000; 2000US-0205515.		
PR	07-JUN-2000; 2000US-0209467.		
PR	28-JUN-2000; 2000US-0214886.		
PR	30-JUN-2000; 2000US-0215135.		
PR	07-JUL-2000; 2000US-0216647.		
PR	07-JUL-2000; 2000US-0216880.		
PR	11-JUL-2000; 2000US-0217487.		
PR	11-JUL-2000; 2000US-0217496.		
PR	14-JUL-2000; 2000US-0218290.		
PR	26-JUL-2000; 2000US-0220963.		
PR	26-JUL-2000; 2000US-0220964.		
PR	14-AUG-2000; 2000US-0224518.		
PR	14-AUG-2000; 2000US-0224519.		
PR	14-AUG-2000; 2000US-0225213.		
PR	14-AUG-2000; 2000US-0225214.		
PR	14-AUG-2000; 2000US-0225266.		
PR	14-AUG-2000; 2000US-0225267.		
PR	14-AUG-2000; 2000US-0225268.		
PR	14-AUG-2000; 2000US-0225270.		
PR	14-AUG-2000; 2000US-0225447.		
PR	14-AUG-2000; 2000US-0225757.		
PR	14-AUG-2000; 2000US-0225758.		
PR	14-AUG-2000; 2000US-0225759.		
PR	18-AUG-2000; 2000US-0226279.		
PR	22-AUG-2000; 2000US-0226681.		
PR	22-AUG-2000; 2000US-0226868.		
PR	22-AUG-2000; 2000US-0227182.		
PR	23-AUG-2000; 2000US-0227009.		
PR	30-AUG-2000; 2000US-0228924.		
PR	01-SEP-2000; 2000US-0229287.		

PR	01-SEP-2000	2000US-0229343	PR	17-NOV-2000	2000US-0249213	PR	17-NOV-2000	2000US-0249213
PR	01-SEP-2000	2000US-0229344	PR	17-NOV-2000	2000US-0249214	PR	17-NOV-2000	2000US-0249214
PR	01-SEP-2000	2000US-0229345	PR	17-NOV-2000	2000US-0249215	PR	17-NOV-2000	2000US-0249215
PR	05-SEP-2000	2000US-0229359	PR	17-NOV-2000	2000US-0249216	PR	17-NOV-2000	2000US-0249216
PR	05-SEP-2000	2000US-0229513	PR	17-NOV-2000	2000US-0249217	PR	17-NOV-2000	2000US-0249217
PR	06-SEP-2000	2000US-0230437	PR	17-NOV-2000	2000US-0249218	PR	17-NOV-2000	2000US-0249218
PR	06-SEP-2000	2000US-0230438	PR	17-NOV-2000	2000US-0249244	PR	17-NOV-2000	2000US-0249244
PR	08-SEP-2000	2000US-0231242	PR	17-NOV-2000	2000US-0249245	PR	17-NOV-2000	2000US-0249245
PR	08-SEP-2000	2000US-0231243	PR	17-NOV-2000	2000US-0249264	PR	17-NOV-2000	2000US-0249264
PR	08-SEP-2000	2000US-0231244	PR	17-NOV-2000	2000US-0249265	PR	17-NOV-2000	2000US-0249265
PR	08-SEP-2000	2000US-0231413	PR	17-NOV-2000	2000US-0249297	PR	17-NOV-2000	2000US-0249297
PR	08-SEP-2000	2000US-0231414	PR	17-NOV-2000	2000US-0249299	PR	17-NOV-2000	2000US-0249299
PR	08-SEP-2000	2000US-0232080	PR	17-NOV-2000	2000US-0249300	PR	17-NOV-2000	2000US-0249300
PR	08-SEP-2000	2000US-0232081	PR	01-DEC-2000	2000US-0250160	PR	01-DEC-2000	2000US-0250160
PR	12-SEP-2000	2000US-0231968	PR	01-DEC-2000	2000US-0250391	PR	01-DEC-2000	2000US-0250391
PR	14-SEP-2000	2000US-0232397	PR	05-DEC-2000	2000US-0251030	PR	05-DEC-2000	2000US-0251030
PR	14-SEP-2000	2000US-0232398	PR	05-DEC-2000	2000US-0251988	PR	05-DEC-2000	2000US-0251988
PR	14-SEP-2000	2000US-0232399	PR	06-DEC-2000	2000US-0256719	PR	06-DEC-2000	2000US-0256719
PR	14-SEP-2000	2000US-0232400	PR	06-DEC-2000	2000US-0251479	PR	06-DEC-2000	2000US-0251479
PR	14-SEP-2000	2000US-0232401	PR	08-DEC-2000	2000US-0251856	PR	08-DEC-2000	2000US-0251856
PR	14-SEP-2000	2000US-0233063	PR	08-DEC-2000	2000US-0251869	PR	08-DEC-2000	2000US-0251869
PR	14-SEP-2000	2000US-0233064	PR	08-DEC-2000	2000US-0251989	PR	08-DEC-2000	2000US-0251989
PR	14-SEP-2000	2000US-0233065	PR	11-DEC-2000	2000US-0251990	PR	11-DEC-2000	2000US-0251990
PR	21-SEP-2000	2000US-0234223	PR	05-JAN-2001	2001US-0259678	PR	05-JAN-2001	2001US-0259678
PR	21-SEP-2000	2000US-0234274	PR			PR		
PR	25-SEP-2000	2000US-0234997	PR			PR		
PR	25-SEP-2000	2000US-0234998	PR			PR		
PR	26-SEP-2000	2000US-0235484	PR			PR		
PR	27-SEP-2000	2000US-0235834	PR			PR		
PR	27-SEP-2000	2000US-0235835	PR			PR		
PR	29-SEP-2000	2000US-0236327	PR			PR		
PR	29-SEP-2000	2000US-0236327	PR			PR		
PR	29-SEP-2000	2000US-0236367	PR			PR		
PR	29-SEP-2000	2000US-0236368	PR			PR		
PR	29-SEP-2000	2000US-0236369	PR			PR		
PR	29-SEP-2000	2000US-0236370	PR			PR		
PR	02-OCT-2000	2000US-0236802	PR			PR		
PR	02-OCT-2000	2000US-0237037	PR			PR		
PR	02-OCT-2000	2000US-0237038	PR			PR		
PR	02-OCT-2000	2000US-0237039	PR			PR		
PR	02-OCT-2000	2000US-0237040	PR			PR		
PR	13-OCT-2000	2000US-0239935	PR					





```

Db 292 TGCTCAATGACATCTCCAGGCTGCAAAATCCCTACTGACACCGCGGACGTCACTGCT 351
Oy 70 alserSerGlyIyArGphearGcYproSerCyArGHisGluValValLeuAspArgH 90
Db 352 CCAATGCTGGAGGCGTTCGCGTCCGCCACACTGCCGCCACAGAGGTGATCATGATGCTC 411
Oy 90 tsGlyValIyTgLy-LeuGlnArGAsn-LeuLeuValGluAsnIleIleAspIleTyrIly 109
Db 412 ACCGAGGTGACCCGTCGACACAGAACCTTGTTGGTGGAGAACATCATGACATCTACAA 471
Oy 109 sGlnGlySerSerArpProLeuHisAlaLysAlaGluGlnHisLeuMetCysGluGluH1 129
Db 472 ACAGAGATGCTTCACAGTGGCGCGCTGCAGAGAGCCACAGCAACCC-ATGAGCAAGAGCA 530
Oy 129 sGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThrcysSerLeuCy 149
Db 531 CGAAGATGAGAAATCAACATCTACTGTCACGTGTGAGGTGCCACCGCTCCATGCTG 590
Oy 149 sLysValPheGlyAlaHisLysAspCysGluValAlaProLeuProThrIleTyrLysAr 169
Db 591 CGAGGTGTTGGATCCACAGAGGCTCGAGGTGGCCCATTTGACAGAGTGTCTTCCAGGG 650
Oy 169 sGlnLysSerGluLeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgValG1 189
Db 651 ACAAAAGACTGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 710
Oy 189 nAlaValIleThrGlnMetGluGluValAlcysGlnThrIleGluAspAsnSerArgArgG1 209
Db 711 GACCATATCATCTACTGAGTGGAGGATTCCTCCGTCGACCAAGAGAAACAGTACACCA 770
Oy 209 nLys-GlnLeuLeuAsnGlnArpPheGluThrLeuCysAlaValLeuGluGlnArpLysG 229
Db 771 AAATAGAGAGCTGAGCCAGAGATTGACACCTGTTATGCTGATGCTGATGATGATGATG 830
Oy 229 LysGluLeuGlnAlaLeuAlaArgGluGlnGluGluLysLeuGlnArpValArgLyl 249
Db 831 GTGAGTGTCTGCTCGCGATCTCGCAAGAGCAGGAAAGAACTTATCTTCTGAGGCGCC 890
Oy 249 euIleArGpGlnTyrGlyAspHisLeuGluGlySerSerLysLeuValGluSerAlaIleG 269
Db 891 TCATCCACGACGTACACGAGGACGCTGGACAGTCCACAGAGCTGGTGGAAACTGCCATCC 950
Oy 269 InsertGluGluProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGluLeuIleAsnL 289
Db 951 AGTCCCTGGAGAGGCTGGGGAGCCACCTTCTCTGACTGCCAACACACTCATCAAAA 1010
Oy 289 ySvalGlyAlaMetSerLysValGluLeuAlaGlyArgProGluProGlyTyrGluSerM 309
Db 1011 GCATTGTGGAAGCTTCCAAAGGCTGCCAGCTGGGGAAGACAGAGCAGGCGCTTTGAGA 1070
Oy 309 etGluGlnPheSerValSerValGluHisValAlaGluMetLeuArgThrIleAspHe- 328
Db 1071 TGGACTCTTTACTTTGATTGAGCATATGACACAGCGCTGAGAGCCATTGACTTTG 1130
Oy 329 -----GlnProGlyA 332
Db 1131 GGACGATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTCCA 1190
Oy 332 laAlaGlyAspGluGlu-AspAspAspMetAlaLeuAsp----- 344
Db 1191 CAGAAAGGAAAGAAAGAAAGACACCACTAAGAGAGCTGATGAATGAGAGGCCCCAGATGC 1250
Oy 345 -----GlyGluGluGlyAsnAlaGlyLeu 352
Db 1251 AGAGAGACTGGAGAGGGTGGGAGGGGCCACAGCGCCTT 1289

```

```

XX DE Human ORFX ORF840 polynucleotide sequence SEQ ID NO:1679.
XX
XX Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
XX vulnary; antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
XX anticonvulsant; osteoplastic; antiarthritic; immunosuppressant; cardiant;
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX hypotensive; dermatologic; immunosuppressive; antiinflammatory;
XX antiviral; antibacterial; antifungal; antirheumatic; antihypertensive;
XX antianemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX cholesterol ester storage; systemic lupus erythematosus; infection;
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX bone damage; cartilage damage; neuroinflammatory disease; coagulation;
XX thrombosis; contraceptive; ss.
XX
XX Homo sapiens.
XX
XX WO200058473-A2.
XX
XX 05-Oct-2000.
XX
XX 31-MAR-2000; 2000WO-US08621.
XX
XX 31-MAR-1999; 99US-0127607.
XX 02-APR-1999; 99US-0127636.
XX 05-APR-1999; 99US-0127728.
XX 30-MAR-2000; 2000US-0540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinketsu RA, Leach M;
XX
XX WPI: 2000-602362/57.
XX P-PSDB: AAB41076.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX
XX Claim 5; Page 1334; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40337 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnary;
XX antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
XX osteoplastic; anticonvulsant; antirheumatic; immunosuppressant;
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatologic; immunosuppressive;
XX antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
XX antihypertensive; antianemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.
XX
XX Sequence 531 BP; 129 A; 143 C; 150 G; 108 T; 1 other;
XX

```

## Alignment Scores:

```

Pred. No.: 9.1e-56 Length: 531
Score: 659.50 Matches: 120
Percent Similarity: 88.16% Conservative: 14
Best Local Similarity: 78.95% Mismatches: 15

```

Query Match: 34,498 Indels: 3  
 DB: 21 Gaps: 2  
 US-09-908-988b-2 (1-366) x AAC75285 (1-531)  
 QY 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluLysGlnLeuIleCysProIleCys 29  
 DB 82 CTGATCCAGAGTGGGAATCCCATGGAGAACTGGAGAGAGAGCTGATCTGCCATCTGC 141  
 QY 30 LeuGluMetPheSerLysProValIleLeuProCysGlnHisAsnLeuCysArgLys 49  
 DB 142 CTGGAGATGTTTACCAAGCAGTGGTCACTTCCCTGCCAGCAACCTGTGCCGAG 201  
 QY 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpIleSerArgLysSerThr 69  
 DB 202 TGTGCCAATGACATCTTCCAGCTGCAGAAATCCCTACTGACACAGCCGGGACCTCA 261  
 QY 70 ValSerSerGlyGlyArgPheArgCysProSerCysArgHisGluValIleAsnAspArg 89  
 DB 262 TCCATGCTCGAGGCGGTTTCCCTGCCCTACCTGCCGACGAGGTGATCATGGATCGT 321  
 QY 90 HisGlyValTyrGlyLeuGlnArgAsnLeuValGlnAsnIleIleAspIleTyrLys 109  
 DB 322 CACGGAGTGTACGGCCTCCAGAGAGAACTGCTGTGGAGAACATCATCGACATCTACAA 381  
 QY 110 GlnGlu---SerSerArgProLeuHisAlaLysAlaGluGlnHisLeuMetCysGluGlu 128  
 DB 382 CAGGAGTGTCTCCAGTGGCGGCGTG-----CAGAAAGGAGTCAACCCATGTACAGAG 435  
 QY 129 HisGlnAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThrCysSerLeu 148  
 DB 436 CACGAAGATGAGAAATCAACATCTACTGCTCACTGCTGAGTGGCCACCTGCTCATG 495  
 QY 149 CysLysValPheGlyAlaHisLysAspCysGluVal 160  
 DB 496 TGCAGCTGTTTGGATCCACAGGCTGCGAGGTG 531  
 RESULT 17  
 ABN78104 ID ABN78104 standard; cDNA; 531 BP.  
 XX  
 AC ABN78104;  
 XX  
 DT 08-JUL-2002 (first entry)  
 XX  
 DE Human Interleukin receptor-like ORF3051 cDNA, SEQ ID NO:6101.  
 XX  
 KM Human; ORF: open reading frame; ORFX: drug screening; diagnosis;  
 KM disease monitoring; cytokine; cell proliferation; cell differentiation;  
 KM immune modulation; haematopoiesis regulation; tissue growth;  
 KM angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;  
 KM thrombolytic; tumour inhibition; bodily characteristic; fertility;  
 KM behaviour; cancer; proliferative disorder; neurological disorder;  
 KM cardiovascular disease; immune system disorder; organ transplantation;  
 KM tissue growth disorder; tissue regeneration disorder; diabetes mellitus;  
 KM hypothyroidism; cholesterol ester storage disease; infection; vulnery;  
 KM vasotrophic; antipsoriatic; antidiabetic; cytoskeletal; neotropic;  
 KM neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;  
 KM cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;  
 KM dermatological; analgesic; virucide; antibacterial; fungicide; gene; ss.

PI Leach MD, Shinkets RA;  
 XX  
 DR WPI: 2002-106200/14.  
 DR P-PSDB; ABP34078.  
 XX  
 PT Novel human polypeptides and polynucleotides useful for diagnosing,  
 PT preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and disorders related to organ  
 PT transplantation  
 XX  
 PS Claim 1; Page 1775; 2508pp; English.  
 XX  
 CC Sequences ABP1028-ABP35561 represent 4534 novel human proteins  
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-  
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses  
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively  
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to  
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX  
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies  
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and  
 CC polypeptides, methods of screening for modulators of ORFX expression or  
 CC activity, and methods of screening individuals for a predisposition to an  
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide  
 CC range of biological activities, such as cytokine, cell proliferation,  
 CC cell differentiation, immune modulation, haematopoiesis regulation,  
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/  
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,  
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,  
 CC and antiinfective activity, and may also be involved in the determination  
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,  
 CC nucleic acids and antibodies may be used in the treatment of cancers,  
 CC other proliferative disorders such as psoriasis and benign tumours,  
 CC neurological disorders such as epilepsy and Alzheimer's disease,  
 CC cardiovascular diseases, immune system disorders, disorders related to  
 CC organ transplantation, disorders of tissue growth and regeneration,  
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester  
 CC storage disease, and infectious diseases caused by viral, bacterial,  
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a  
 CC source of primers and probes, in the detection of ORFX genomic sequences  
 CC or transcripts, in the identification and cloning of homologous  
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX  
 CC nucleic acids may additionally be used to produce transgenic animals  
 CC which may be useful for studying the function and/or activity of ORFX  
 CC protein, and in drug screening. The ORFX proteins may also be used as  
 CC immunogens to generate specific antibodies, which are useful in the  
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.  
 XX  
 SQ Sequence 531 BP; 129 A; 143 C; 150 G; 108 T; 1 other;  
 Alignment Scores:  
 Pred. No.: 9,1e-56 Length: 531  
 Score: 659.50 Matches: 120  
 Percent Similarity: 88.16% Conservative: 14  
 Best Local Similarity: 78.95% Mismatches: 15  
 Query Match: 34,498 Indels: 3  
 DB: 24 Gaps: 2  
 US-09-908-988b-2 (1-366) x ABN78104 (1-531)  
 QY 10 LeuLeuGlyAspAlaHisAsnMetAspAsnLeuGluLysGlnLeuIleCysProIleCys 29  
 DB 82 CTGATCCAGAGTGGGAATCCCATGGAGAACTGGAGAGAGCTGATCTGCCATCTGC 141  
 QY 30 LeuGluMetPheSerLysProValIleLeuProCysGlnHisAsnLeuCysArgLys 49  
 DB 142 CTGGAGATGTTTACCAAGCAGTGGTCACTTCCCTGCCAGCAACCTGTGCCGAG 201  
 QY 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpIleSerArgLysSerThr 69  
 DB 202 TGTGCCAATGACATCTTCCAGCTGCAGAAATCCCTACTGACACAGCCGGGACCTCA 261  
 QY 70 ValSerSerGlyGlyArgPheArgCysProSerCysArgHisGluValIleAsnAspArg 89

Db 262 TCCATGTCGTGAGGCCGTTTCCGCTGCGCTACTACGCGCCGACGAGGTGATCATGATGTCGT 321  
Qy 90 HtSGlyValTyrGlyLeuGlnArgAsnLeuValAluAsnIleIleAspIleTyrLys 109  
Db 322 CACGAGGTGTACGGCTGACGAGAACCTGCTGTGAGAACATCATCATCATCAAA 381  
Qy 110 GlnGlu---SerSerArgProLeuHisAlaLysAlaGlnGlnIleSleuMeCysGlnGlu 128  
Db 382 CAGGAGTGTCTCCAGTGTGCGCCGCTG-----CAGAGGGGAGCTCACCCCATCTACAGAGAG 435  
Qy 129 HisGluAspGluLysIleAsnIleTyrCysLeuSerCysGluValProThrCysSerLeu 148  
Db 436 CACGAAAGATGAGAAATCAACATCTACTGTCTCAGCTGTGAGGTGCCACCTGCTCATG 495  
Qy 149 CysLysValPheGlyAlaHisLysAspCysGluVal 160  
Db 496 TGCAGAGTGTGTGGATCCACAGAGGCTGCGAGGTG 531  
RESULT 18  
AAS26314  
ID AAS26314 standard; cDNA; 587 BP.  
AC AAS26314;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human cDNA encoding a novel secreted protein, Seq ID 493.  
XX  
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic;  
KW cytosolic; cardiant; vasotopic; cerebroprotective; noctropic;  
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
KW vulnerrary; secreted protein; rheumatoid arthritis;  
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;  
KW cerebrovascular disorder; cerebral ischaemia; angioneuosis;  
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;  
KW corneal infection; wound healing; epithelial cell proliferation;  
KW skin ageing; food additive; preservative; antiproliferative.  
XX  
OS Homo sapiens.  
XX  
PN WO200155322-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01341.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 30-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229309.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234224.  
PR 25-SEP-2000; 2000US-0234597.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.



PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX  
XX WPI: 2001-468783/53.  
DR  
XX  
XX P-PSDB: AA016327.  
PT  
PT New nucleic acid molecules encoding 461 human secreted proteins for  
PT diagnosing, preventing, treating or ameliorating medical conditions and  
PT used as food additives or preservatives -  
PS  
PS Claim 1; SEQ ID NO 493; 980bp; English.  
XX  
XX  
XX The invention relates to isolated nucleic acid molecules and their  
CC encoded secreted proteins. The nucleic acids and proteins are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They  
CC are also used in diagnosing a pathological condition or susceptibility  
CC to a pathological condition. Antibodies to the proteins can also  
CC be used in alleviating symptoms associated with the disorders and in  
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
CC immunoassay assays (ELISA). Disorders which are diagnosed or treated  
CC include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi  
CC and ocular disorders e.g. corneal infection, and many other  
CC disorders listed in the specification. The polypeptides can also  
CC be used to aid wound healing and epithelial cell proliferation, to  
CC prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capability, fat content, lipid, protein, carbohydrate, vitamins,  
CC minerals, cofactors and other nutritional components. The present

CC sequence encodes a novel secreted protein of the invention.  
Alignment Scores:  
Pred. No.: 3,14e-51 Length: 587  
Score: 614.00 Matches: 115  
Percent Similarity: 86.00% Conservative: 14  
Best Local Similarity: 76.67% Mismatches: 19  
Query Match: 32.11% Indels: 3  
DB: 22 Gaps: 1  
US-09-908-988b-2 (1-366) x AAS26314 (1-587)  
QY 4 ThrValGlyPheLeuProLeuGluAlaAspAlaHisAsnMetAspAsnLeuGluVal 23  
DB 145 TCTGGAATTACAAATCTTTTCCAAAGAGACAGACCATGGATTAACCTTAGAGAAACAA 204  
QY 24 LeuIleCysProIleCysLeuGluMetPheSerLeuProValValIleLeuProCysGln 43  
DB 205 CTCATCTGCCATCTGCTAGAGATGTCACGAAACCTGTGTGATTCCTCCTGTGAC 264  
QY 44 HisAsnLeuCysArgLeuCysAlaAsnAspValPheGlnAlaSerAsnProLeuTrpGln 63  
DB 265 CACACCTGTGTAGGAATGTCACCATGATATTTTCCAGGCTCTAACCCGTATTGCCC 324  
QY 64 SerArgGlySerThrThrValSerSerGlyGlyArgPheArgCysProSerCysArgHis 83  
DB 325 ACAAGAGAGATGACACCATGGCATCAGGGGCGCATTCGCGTCCCATCTGTAGACAT 384  
QY 84 GluValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuValGluAsn 103  
DB 385 GAAGTGGTTTGGATGATGACATGGGTATATGACTTCAGAGAACTGCTGGTGGAAAT 444  
QY 104 IleIleAspIleTyrLeuGlnGluSerSerArgProLeuHisAlaValGluGlnHis 123  
DB 445 ATCATTTGACATCTACAGACGAGAGTCCACAGGCCA-----GAAGAATAATCGACAC 498  
QY 124 LeuMetCysGluGluHisGluAspGluValIleAsnIleTyrCysLeuSerCysGluVal 143  
DB 499 CCCATGTGGGAGAACATGATGAGAGAGCGCATCATCTACTGTGTGACTGGAGATA- 557  
QY 144 ProThrCysSerLeuCysLysValPheGly 153  
DB 558 CCCACCTGCTCTGTGCAAGGTGTTGTT 587  
RESULT 19  
ABV21919/C  
ID ABV21919 standard; cDNA; 4345 BP.  
XX  
XX  
XX AC ABV21919;  
XX  
XX  
XX DT 13-SEP-2002 (first entry)  
XX  
XX DE Human prostate expression marker cDNA 21910.  
XX  
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX pharmacogenomic marker; gene; ss.  
XX  
XX OS Homo sapiens.  
XX  
XX  
XX PN WO200160860-A2.  
XX  
XX PD 23-AUG-2001.  
XX  
XX  
XX PF 20-FEB-2001; 2001WO-US05171.  
XX  
XX  
XX PR 17-FEB-2000; 2000US-183319P.  
XX PR 16-MAR-2000; 2000US-189862P.  
XX PR 25-MAY-2000; 2000US-207454P.  
XX PR 09-JUN-2000; 2000US-211314P.  
XX PR 18-JUL-2000; 2000US-219007P.  
XX PR 13-DEC-2000; 2000US-255281P.  
XX  
XX  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.



XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX Schlegel R, Endege WO, Monahan JE;  
 XX WPI: 2001-662795/7.6.  
 XX  
 XX Novel isolated nucleic acid molecule associated with cancerous state of  
 XX prostate cells and correlating with presence of prostate cancer, useful  
 XX for detecting presence of prostate cancer, stage of prostate cancer  
 XX  
 XX Claim 1: Page 3836-3837; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising  
 a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 specification or its complement. (I) is useful for:  
 (a) assessing whether a patient is afflicted with prostate cancer;  
 (b) monitoring the progression of prostate cancer in a patient;  
 (c) assessing the efficacy of a test compound to inhibit prostate  
 cancer in a patient;  
 (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
 in a patient;  
 (e) selecting a composition for inhibiting prostate cancer in a patient;  
 (f) assessing the prostate cell carcinogenic potential of a compound;  
 (g) determining whether prostate cancer has metastasized in a patient;  
 (h) assessing the aggressiveness or indolence of prostate cancer in a  
 patient;  
 (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

Sequence 4345 BP: 1321 A; 874 C; 801 G; 1319 T; 30 other;

Alignment Scores:  
 Pred. No.: 1,12e-31 Length: 4345  
 Score: 427.50 Matches: 112  
 Percent Similarity: 44.30% Conservative: 63  
 Best Local Similarity: 28.35% Mismatches: 127  
 Query Match: 22.36% Indels: 93  
 Gaps: 10

US-09-908-988B-2 (1-366) x ABV22250 (1-4345)

OY 19 AsnLeuGluValGlnLeuIleuLeuProIleuGlnLeuIleuMetPheSerIleuVal 38  
 DB 3882 AATATCGAAAGGAGCTCTTTGCCCCAGCATGCAAGAGCTGTTTACCACCA--TTG 3826  
 OY 39 IleuProGlnHisAsnLeuGlnValGlnValGlnValGlnValGlnValGlnVal 51  
 DB 3825 ATTCTGCTTGGCAACATGATGATCTGATTAATGTGTAAGAAGCTCTGCTGCTC 3766  
 OY 52 -----AsnAspVal----- 54  
 DB 3765 GATGATTCATTCACGATGCTGGATCAGACAACTCCAAATCAAGCAGTCTCGACTCGG 3706  
 OY 55 -----PheGlnAlaSerAsnProLeuIleuValGlnValGlnValGlnValGlnVal 65  
 DB 3705 CTCCTCCCTCCCTAGTATGATTAATTTGACCAATTTACAGACAGCTGGAACCGCAAT 3646  
 OY 66 Gly-----SerThrThrValSerSerGlyGlyArgPheArgCysProSerCysArg 82  
 DB 3645 TCATTCAGCCCGAGGACACTGTT-----TTCCCTTCCCTGGCTGTGAG 3601  
 OY 83 HisGluValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeuValGlu 102  
 DB 3600 CATGATGTGGATCTTGGAGACGAGGAATCAATGCTGTGTTGAAACTTCACCTTGAA 3541  
 OY 103 AsnIleIleAspIleTyrIle----- 109  
 DB 3540 ACTATTGTGGAAAGATATGCTCAAGCAGCTAGGGCAGCCACAGCCATTATGTGACCTT 3481  
 OY 110 -----GlnGluSerSerArg----- 114  
 DB 3480 TGTAAACCAACACCTCAAAATTCACAAAAGCTGATGATGATGCAAGTTACTGC 3421

OY 115 -----ProLeuHisAlaValGlnGlnHis----- 123  
 DB 3420 AATGATGCTTCAAAATTCATCACCCTTGGGGTACATATAAAGCTCAAGTATGATGT 3361  
 OY 124 -----LeuMetCysGluGlnHisGluAspGlu 133  
 DB 3360 GGTCCAACTACTAATTCAGACCCAAAGATTTTATATGTGCCCCAGACATGAAACAGAGA 3301  
 OY 134 IleAsnIleTyrCysLeuSerCysGluValProThrCysSerIleCysValPheGly 153  
 DB 3300 ATTAACATGATCTGCAATTAATGATGAGAGCCGTTTCCATCTGTGTATGGGTGT 3241  
 OY 154 AlaHisIleAspCysGluValAlaProLeuProThrIleTyrIleAspArgGlnIleSerGlu 173  
 DB 3240 AATCATGCCAACCCGCTTAACCATATGAGCAGCGCTACAAAACCTTAAGGAAAG 3181  
 OY 174 LeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgValGlnAlaValIleThr 193  
 DB 3180 CTTTCAAGAGATATGATTAATGATTAAGGAAAGCCAGCTGAAGATCAATATCT 3121  
 OY 194 GlnMetGluValCysGlnThrIleGluAspAsnSerIleArgGlnIleGlnLeu 213  
 DB 3120 GAACTAACTTGTATGTAAGAAACAGAGTGAATGAGAGAGGCTTAAGAAAGCA 3061  
 OY 214 AsnGlnArgPheGluThrLeuCysAlaValLeuGluGlnArgGlyGlnLeuGln 233  
 DB 3060 ATTACACATTTTGAAGAAGCTTTGGAAGTTTGGAGAGAGAAATCATCTGTTTGA 3001  
 OY 234 AlaLeuAlaArgGluGlnGluGlnIleuValGlnValArgGlyLeuIleArgGlnTyr 253  
 DB 3000 GCAATTCAGCTCTTAAGAAACTAAGATTAAGCAAAATTCAGACTCAAAATGGAAGATAC 2941  
 OY 254 GlyAspHisLeuGluGlnIleuValGlnValGlnValGlnValGlnValGlnValGlnVal 273  
 DB 2940 CAGGACTCTTGAAGACAAATGA---CTTGTGGATATGCTCAAGAGCTTAAGAGAG 2884  
 OY 274 ProGlnMetAlaLeuTyrLeuGlnAlaValGlnValGlnValGlnValGlnValGlnVal 293  
 DB 2883 ACAGATCAGTCTTGTGTTGTCAGACAGCAAGCACTCCACTCAGAAATACAGAAAGCC 2824  
 OY 294 SerIleValGlnLeuAlaGlyArgProGluProGlyTyrGlnSerMetGluIlePheSer 313  
 DB 2823 ACAGATCTTTGAAGAGCTTTAGACCT--GCAGCTCAGACTCTTTTGAAGACTATGTT 2767  
 OY 314 ValSerValGlnHisValAlaGlnMetLeuArgThrIleAspPhe 328  
 DB 2766 GTTAATACCTTAACAAACAGAACTTCTTGGAGAAATTAATCTTT 2722

Search completed: December 3, 2002, 14:10:55  
 Job time : 299 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 3, 2002, 14:05:55 ; Search time 2707 Seconds

(without alignments)  
3934.846 Million cell updates/sec

Title: US-09-908-988B-2

Sequence: 1 MNFTVGFKPLGDHNNMDNL.....EGNAGLEERLDEVEGSLH 366

Scoring table: BLAST62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 1451402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-O=cg2\_1/USPFO.spool/US09080888/runat\_26112002.112154.11224/app.query.fasta.1.519  
-MODEL=frame+ p2n.model -DEV=xlh  
-DB=GenBdl -QPM=fastap -SUFFIX=ige -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human4.0.cdi -LIST=45  
-DOCALLIGN=200 -THS\_SCORE=pcpt -THR\_MAX=100 -THR\_MIN=0 -ALIGN=20 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09080888 @CGN.1.1.1616 @runat.26112002.112154.11224 -NCPU=6 -ICPU=3  
-NO\_XLPRXY -NO\_MMAR -LARGESOURCE -NEG\_SCORES=0 -WATT -LONGLOG -DEV\_TIMECUT=120  
-WARN\_TIMECUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pt:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_iny:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1912	100.0	1431	6	AX418848	AX418848 Sequence
2	1912	100.0	1448	10	AF294790	AF294790 Mus muscu
3	1743	91.2	1329	9	HS291714	AJ291714 Homo sapi
4	1261.5	66.0	1500	6	AX060632	AX060632 Sequence
5	1162	60.7	1750	6	AX418850	AX418850 Sequence
6	1161	60.7	1750	6	BC007750	BC007750 Homo sapi
7	1161	60.7	1925	9	AK091728	AK091728 Homo sapi
8	1161	60.7	2202	9	HS291712	AJ291712 Homo sapi
9	1161	60.7	2634	9	AK091310	AK091310 Homo sapi
10	1157	60.5	1810	9	HS243488	AJ243488 Homo sapi
11	1157	60.5	2098	9	HS243489	AJ243489 Homo sapi
12	1119.5	58.6	1211	9	BC015717	BC015717 Homo sapi
13	1110	58.1	1397	6	AX418852	AX418852 Sequence
14	1109	58.0	1861	10	AY059627	AY059627 Rattus no
15	1102.5	57.7	2097	9	HS291713	AJ291713 Homo sapi
16	1098.5	57.5	1756	9	HS276484	AJ276484 Human mRN
17	1098.5	57.5	1764	6	AX274927	AX274927 Sequence
18	1098.5	57.5	1764	9	AK056942	AK056942 Homo sapi
19	1098.5	57.5	1764	9	AK053673	AK053673 Homo sapi
20	866	45.3	1746	9	AB047601	AB047601 Macaca fa
21	659.5	34.5	531	6	AX313116	AX313116 Sequence
22	624	32.6	199296	2	AC114619	AC114619 Mus muscu
23	624	32.6	209211	2	AC084883	AC084883 Mus muscu
24	624	32.6	216193	10	AC109608	AC109608 Mus muscu
25	601	31.4	1903	9	AF361946	AF361946 Homo sapi
26	594.5	25.9	118847	2	AC013413	AC013413 Homo sapi
27	494.5	25.9	164772	2	AC095184	AC095184 Rattus no
28	439	23.0	1733	9	AF041206	AF041206 Homo sapi
29	439	23.0	2262	10	AF186461	AF186461 Rattus no
30	437	22.9	3334	9	AF035360	AF035360 Homo sapi
31	437	22.9	3352	9	AF230977	AF230977 Homo sapi
32	437	22.9	3422	9	AF041209	AF041209 Homo sapi
33	437	22.9	3452	9	AF230976	AF230976 Homo sapi
34	437	22.9	3452	9	HSXPRE	Y13667 Homo sapien
35	437	22.9	3575	9	AF269101	AF269101 Homo sapi
36	437	22.9	3628	9	AF041208	AF041208 Homo sapi
37	436	22.8	2670	10	MMY14848	Y14848 Mus muscu
38	432	22.6	2004	5	AF269102	AF269102 Gallus ga
39	430	22.5	2511	10	AF026565	AF026565 Mus muscu
40	428	22.4	2298	5	AF374463	AF374463 Gallus ga
41	427.5	22.4	4074	10	HS272269	AJ272269 Homo sapi
42	427	22.3	3348	10	AF186460	AF186460 Mus spret
43	425.5	22.3	2716	9	AB056399	AB056399 Macaca fa
44	424.5	22.2	1875	9	AK026882	AK026882 Homo sapi
45	421.5	22.0	2679	9	AB063010	AB063010 Macaca fa

RESULT 1

#### ALIGNMENTS

AX418848  
LOCUS AX418848 1431 bp DNA linear PAT 18-JUN-2002  
DEFINITION Sequence 1 from Patent WO0206318.  
ACCESSION AX418848  
VERSION AX418848.1 GI:21523712  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 Olson, E.N. and Spencer, J.A.  
METHODS and compositions for stabilizing microtubules and intermediate filaments in striated muscle cells  
Patent: WO 0206318-A 1 24-JAN-2002.  
JOURNAL BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)  
FEATURES  
source Location/Qualifiers  
1.1431  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
199.1299  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD35455.1"  
/db\_xref="GI:21523713"  
/translation="MNFTVGFKPLDGDHNMNLEKOLICPLEMEFSPVILPCOH  
NLCKRCAVDFOASNPIMOSRGSTSSGSRFCRSHVYLDHGVGLONLVE  
NLIDTYKROESRPLHAKADHLMCEHDEPEKINYLSCSEVPSCICKVGAIKDCV  
APLEPTIKROKSELSDGIMLVAGNDROAVITOMEVOCOTIDNSRQQLNORE  
TLCAVLERKEKLQALAREDEKLQAVNGLRQVGLHLEGGSKLVESAQSMEEDPM  
ALYLOQAKELINKVAMSKVELAGREPECESEMEQFSVEHVAEMLRITDFOPGAAG  
DEEDDMALDEEGNAGLEERLDYVEGSGH"  
BASE COUNT 338 a 384 c 473 g 236 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.46e-146 Length: 1431  
Score: 1912.00 Matches: 366  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
US-09-908-988b-2 (1-366) x AX418848 (1-1431)  
QY 1 Metasphehrhva1g1yphelysProleu1eug1yaspAlah1sasnmetasp1e1eu 20  
Db 199 ATGAACCTTCAGGGTGGTTTCAAGCCGCTGCTAGGGGATGCGCACACATGACAACTTG 258  
QY 21 Glulysgl1nleu1le1ecysPro1le1ecysleu1eug1metPheSer1yProVal1le1eu 40  
Db 259 GAGAGAGCGATCATTTGGCCCATCTGCTGAGATGTCTCCAAAGCCCGTGGATCTTG 318  
QY 41 Procyg1n1h1sasn1eucysArg1yScysAla1asn1aspVal1Phe1n1Ala1Ser1nPro 60  
Db 319 CCGTGCACACCAACCTGTGCGCAAGTGTGCCAAGAGCTTCCAGGCGCTTAATCCT 378  
QY 61 Leutrp1n1Ser1Arg1y1Ser1Thr1hVal1Ser1Ser1y1Gly1yArgPhe1yGysPro1Ser 80  
Db 379 CTGGGCAATCCCGGGGCTCCCAACAGGCTGTCTTCAGGAGGACGTTCCGATGCCATCT 438  
QY 81 CysArg1h1sg1uVal1Val1leu1aspArg1h1sg1yVal1Tyrg1y1eug1n1Arg1asn1leu 100  
Db 439 TGTAGGCGACGAGTGTCTCGTGCACAGGCATGTGCTATAGCGCTGCAGCGAAGCTGCTA 498  
QY 101 Val1glu1asn1le1le1asp1le1Tyrg1y1sg1n1glu1Ser1Ser1y1Pro1leu1h1sAla1y1sAla 120  
Db 499 GTGGAGAACATCATTTGACATCTACAAAGCAGGAGTCTCCCGGCGCACTGCACGCGCAAGCT 558  
QY 121 Glu1n1h1s1eumetCys1glu1n1h1sg1u1asp1glu1y1le1asn1le1Tyrg1y1Ser1e1uSer 140  
Db 559 GAACAGCACCTCATGTGTGTGAGGAGCACAGAGCAGAGAAATCAACATCTACTGCTGAGC 618

QY 141 Cysgl1uVal1ProthrcysSer1eucys1yVal1Phe1yAla1h1s1y1sAsp1cysgl1uVal 160  
Db 619 TCGAGAGTGGCCACACTCTCTCTCTGCAAGGTTTTCGGCCGCCACAGAGACTGTACAGTG 678  
QY 161 AlaPro1euprothrl1eTyrg1y1sArg1n1y1sSer1e1u1e1uSer1asp1y1le1Ala1met 180  
Db 679 GCCCTCTGCCACCATTTACAAACGCCAGAAAGATGAGTGAGGATGCGCATCCGATG 738  
QY 181 LeuVal1Ala1gl1ysn1aspArg1yAla1Val1le1Thrg1n1met1glu1Val1y1sg1n 200  
Db 739 CTGTGGCGGGCAATGACCTGTGTACAGCACTGATCACCCAGATGAGAGAGGTGTGCCAG 798  
QY 201 Thr1l1glu1asp1asn1er1Arg1n1y1sg1n1leu1euan1gn1n1Arg1Phe1n1h1eu 220  
Db 799 ACATTGAGACACAAACAGCCCGACAGACAGAACTGTTAAACAGAGTTCCAGACCTG 858  
QY 221 CysAla1Val1eug1n1glu1Arg1y1Gly1leu1eug1n1Ala1leu1Ala1Arg1n1glu1n 240  
Db 859 TCGCGGTTTTGGAGAGCGCAAGGCGCAACTGCTTCMAAGCACTGCGCGGAGCAGAG 918  
QY 241 Glulysleu1n1Arg1yVal1Arg1y1leu1le1Arg1n1Tyrg1y1Asp1h1s1eug1n1y1Ser 260  
Db 919 GAGAAATTTGCAGCGCGTGTGCGGGCTTCATCCGACGATACGAGACACTTGGAGGGCTTC 978  
QY 261 Ser1y1e1uVal1glu1Ser1Ala1le1gl1n1Ser1Met1glu1n1Pro1gl1n1Met1Ala1e1uTyrg1eu 280  
Db 979 TCAAACTGCTGAGATGCCGCACTTCATGAGAGAGCCGCAATGCTCTTACTTC 1038  
QY 281 Glu1n1Ala1y1sg1n1leu1le1asn1y1sVal1Gly1Ala1Met1Ser1y1sVal1glu1eua1la1y 300  
Db 1039 CAGCAGCAAGAGAGCTGATCAACAAGCTGCGGCGCAATGTCGAAGGTGAGAGCTGACAGA 1098  
QY 301 ArgPro1glu1Pro1Tyrg1y1r1g1u1Ser1Met1glu1n1Phe1Ser1Val1glu1h1s1Ala1 320  
Db 1099 CGCCCGAGCGAGCTATGAGACATGAGCAATCTCTGTGACCGTGTGACACACTGCGC 1158  
QY 321 Glu1Met1eua1r1g1h1le1a1r1Phe1n1Pro1gl1y1Ala1a1g1y1asp1glu1n1asp1asp 340  
Db 1159 GAAATTTTGCAGACATTCGACTTCCACCGCGCGCCCTGGGATGAAGAGATATACGAC 1218  
QY 341 MetAla1eua1s1p1y1glu1glu1y1asn1Ala1gl1y1eug1n1glu1n1Arg1eua1s1p1y1Pro 360  
Db 1219 ATGCGTTTGGATGGGAGGAGGCGCAATGCGGCGCTGAGAGAGAGCGGCTGACAGTGCAC 1278  
QY 361 Glu1y1Ser1y1e1u1h1s 366  
Db 1279 GAAGCTCAGGCTTGAC 1296  
RESULT 2  
AF294790 1448 bp mRNA linear ROD 08-SEP-2000  
LOCUS AF294790  
DEFINITION Mus musculus RING-finger protein MURF mRNA, complete cds.  
ACCESSION AF294790  
VERSION AF294790.1 GI:9945009  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 Spencer, J.A., Eliazzer, S., Ilaria, R.L. Jr., Richardson, J.A. and Olson, E.N.  
TITLES Regulation of microtubule dynamics and myogenic differentiation by MURF, a striated muscle RING-finger protein  
JOURNAL J. Cell Biol. 150 (4), 771-784 (2000)  
MEDLINE 20411220  
PUBMED 10953002  
REFERENCE  
2 (bases 1 to 1448)  
AUTHORS Spencer, J.A. and Olson, E.N.  
TITLES Direct Submission  
JOURNAL Submitted (09-AUG-2000) Molecular Biology, UTSW, 6000 Harry Hines Blvd., Dallas, TX 75390-9148, USA  
FEATURES  
Location/Qualifiers

source

1. 1448  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/tissue\_type="cardiac muscle"  
199. 1299  
/codon\_start=1  
/product="RING-finger protein MURF"  
/protein\_id="AA03076.1"  
/db\_xref="GI:9945010"  
/translation="MNFVGFKPLIGDAHNMNDLEKOLICPTCLEMSKPYVILPCOH  
NLCRKANDYFQASNPIMOSRSTPSSGFRSCSREVLVDHGYGLQKLLVE  
NIIDIKYQESSRPLHAKAEQHLKCEBHEDEINITYCLSEVPTCLCVFGAKHCEV  
APLPTITKROKSELSDGILMLVAGNDROAVITQMEVCOITEDNSRQKOLLNRF  
TLCAVEERKGLLQALAREQEKLOVRVGLIROYGDHLESSKLVESAIOSEEPOM  
ALYLOAKELINKVAMSKVELAGRPPEYSEMEQFSVSEHVAEMLRITIDPQAG  
DEEDDMALDGEENAGLEERLDVPEGSGLH"

BASE COUNT 355 a 384 c 473 g 236 t

ORIGIN

Alignment Scores:

Pred. No.: 2.5e-146 Length: 1448  
Score: 1912.00 Matches: 366  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-09-908-988b-2 (1-366) x AF294790 (1-1448)

QY 1 Metanpherhrrvaliglyphelphylproleuleuglyaspaalahaismetaspasleu 20  
DB 199 ATGACTTCACGGTGGGTTTCAACCGCTGCTAGGGGATGGCCACATGACAACTTG 258

QY 21 Glubysgluleuilecysproillecysleuglumetphaserlysprovalyalleu 40  
DB 259 GAGAGACAGCATTTGCCCATCTGCTGAGAGATGTTCTCCAAACCCGTGTGATCTTG 318

QY 41 Procyglinhiasnleucysarglyscysalaasnapvalpheglinalaseranpro 60  
DB 319 CCGTGCACACAACTGTGCGGCAAGTGTCCAAAGCATCTTCCAGCCCTCTATCT 378

QY 61 Leutrpinserrarglyserthrthrvalserserglyyargpheargcyproser 80  
DB 379 CTGTGCAATCCCGGGGCTCCACAAAGGTGTCTTCAAGAGACGTTTCCGATGCCATCT 438

QY 81 Cysarghislgluvalleuasparghislgluvaltyrlyleuglinalarqasleu 100  
DB 439 TGTAGGACAGAGGTGTCTCGAGACAGCATGTGTCTATGGCTCTCAGCGAACCCTGTA 498

QY 101 ValgluasnlleileasplietrlylsglngluserSerargproleuinhialalsala 120  
DB 499 GTGGAGAACATCATTTGACATCTACACAGAGATCTCCGGCCACTCAGCCCAAGGCT 558

QY 121 Gluglnhislseumetcyssgluglulhislgluaspglulyslleasplietrlycysleuser 140  
DB 559 GAACAGACACTCATGTGTGAGAGACAGAGACGAGAGATCAATCTACTGCTGAGC 618

QY 141 Cysgluvalprothrlycysserleucyslyvalpheglialhisllyaspcysgluval 160  
DB 619 TCGAGAGTGGCCACTCTCTCTGCAAGGTTTTCGGCGCCCAAGAGACTGTAGGTG 678

QY 161 AlaProleuprothrlietrrlylrsargglulyssergluleuseraspglytlealmet 180  
DB 679 GCCCTCTGCGCCACCATTTACAAACGCCAGAGAGTGTGAGCAGATGATCGCATG 738

QY 181 Leuvalalagllysnasparvalglinalavallierphrglnmetglulvalcysgln 200  
DB 739 CTGTGTGGCGGCAATGACCTGTGTGACAGCATGATCCAGATGAGAGAGTGTGCGAG 798

QY 201 Thillegluaspanserargargglulysgluleuansnglnarphegluthrleu 220  
DB 799 ACCATTGAGAGCAAGCCGACAGACAGAACTGTTAAACGAGAGGTTGAGACCCCTG 858

QY 221 Cysalavalleuglularglylsgluleuenglinalaleualarglulglnu 240  
DB 859 TGGCCGGTTTGGAGAGACCGCAAGGCGAAGCTGTTCAAGCATGGCCCGGAGAGAG 918

QY 241 Glubysleuglnarvalarglylleuilearqglntyrlyaspshisdeuglulysr 260  
DB 919 GAGAAATTGACGCCGTCGGGGCTCATCCGCCAGTACGGAGACCACTTGGAGGGCTCC 978

QY 261 Serlysluvalgluseralalleinserrmetgluluproglnmetalaleturleu 280  
DB 979 TCAAGGTGTGGATCCGCTCATCCAGTCCATGAGAGAGCCGCAATGAGCTCTTACCTC 1038

QY 281 Glnglnalavysgluleuileasnlysalgllyalmetserlyvalgluenuagly 300  
DB 1039 CAGAGCGCAAGAGCTGATCAACAAGGTCCGGGCAATGTGCAAGCTGAGCTGCAGGA 1098

QY 301 Argpogluproglytyrlyglusermetglulnpheservalservalglulhisvala 320  
DB 1099 CGGCGGAGCCAGGCTATGAGACATGAGCAATCTCTGTGAGCGTGGACAGCTGGCC 1158

QY 321 Glumetleuargrthrlleasppheglinprogllyalalaglyaspglulaspasp 340  
DB 1159 GAAATGTTGCAACCATGACTTCCAGCGGGCCCGCTGGGAGTGAAGAGATGACGAC 1218

QY 341 Metalaenuaspglyglulglulgllyasnalaglyleuglulglulargleuaspa 360  
DB 1219 ATGGCTTGGATGGGAGAGAGAGGCAATCGGGGCTGTGAGAGAGAGAGGCTGACCTGCCA 1278

QY 361 GlulyserrglyleuHis 366  
DB 1279 GAAAGCTCAGGCGCTGCAC 1296

RESULT 3  
HSA291714 HSA291714 1329 bp mRNA linear PRI 06-JUL-2001  
LOCUS Homo sapiens mRNA for RNF30 gene for ring finger protein 30.  
DEFINITION AJ291714  
VERSION AJ291714.1 GI:13160387  
KEYWORDS MURF-3 gene; muscle specific ring finger protein.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1329)  
REFERENCE  
AUTHORS Centner,T., Yano,J., Kimura,E., McElhinny,A.S., Pelin,K.,  
Witt,C.C., Bang,M.L., Tromblais,K., Granzler,H., Gregorio,C.C.,  
Sortimachi,H. and Labell,S.  
TITLE Identification of muscle specific ring finger proteins as potential  
regulators of the titin kinase domain  
JOURNAL J. Mol. Biol. 306 (4), 717-726 (2001)  
MEDLINE 21140110  
PUBMED 11243782  
REFERENCE 2 (bases 1 to 1329)  
AUTHORS Centner,T.  
TITLE Direct Submission  
SUBMITTED (12-FEB-2001) Centner T., Structure and Biocomputing,  
EMBL-Heidelberg, Meyerhofstrasse 1, Heidelberg 69112, Germany  
REMARK Revised by author 22-FEB-2001  
FEATURES  
source  
1..1329  
Location/Qualifiers  
1..1329  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1..3  
/note="putative alternative translation initiation codon"  
/evidence=not\_experimental  
49..1203  
/gene="RNF30"  
49..1203  
/gene="RNF30"  
/note="RNF30"  
/codon\_start=1  
/product="ring finger protein 30"  
/protein\_id="CAC32842.1"

```

/db_xref="GI:13160389"
/translation="MDNLEKOLICPLEEMFSKPVVILPCOHNLCRKCANVFOASNP
LMOSSGSTVSSGFRPCSCRHVEVLDHRHVGGLQNRNLVENIIDIKQESSRPLHS
KAOHLMCEHEERKINICYLSCVEPCSLCKVFAGHDCVEAPLPTIKRKODLT
LPLRECSNTNTYCSLDLPSSDPPIIASONTKIDSELSPILAMLVAGNRVQVIT
TOMEVCOCTEDNSRROKOLITOREPESICAVLEERKGLLOLAGEERKLOVRGLI
ROYGDHLEASSKIVESAIOSMEBOMALYIOAKELINKVMSRVELAGREPES
MEQFIVREHVAEMLRITDFOPASGEEVEVAPDEESAGFEERPDGP"
join(49..515,642..1203)
/gene="RNF30"
/note="short isoform"
/codon_start=1
/product="ring finger protein 30"
/protein_id="CAC32841.1"
/db_xref="GI:13160388"
/translation="MDNLEKOLICPLEEMFSKPVVILPCOHNLCRKCANVFOASNP
LMOSSGSTVSSGFRPCSCRHVEVLDHRHVGGLQNRNLVENIIDIKQESSRPLHS
KAOHLMCEHEERKINICYLSCVEPCSLCKVFAGHDCVEAPLPTIKRKODLT
GIAMLVAGNRVQVITOMEVCOCTEDNSRROKOLITOREPESICAVLEERKGLLOL
LAGEERKLOVRGLIROYGDHLEASSKIVESAIOSMEBOMALYIOAKELINKYGA
MSRVELAGREPESMEQFIVREHVAEMLRITDFOPASGEEVEVAPDEESAGP
FEERPDGP"
BASE COUNT      310 a      375 c      414 g      230 t
ORIGIN
Alignment Scores:
Pred. No.:      1,2e-132      Length:      1329
Score:          1743.00      Matches:      342
Percent Similarity: 87.13%      Conservative: 10
Best Local Similarity: 84.65%      Mismatches: 8
Query Match:     91.16%      Indels:      44
DB:              Gaps:      2
US-09-908-988b-2 (1-366) x HSA291714 (1-1329)
QY      1      MetaspherthalglyphelyProleuLeuGlyAspAlaHisAsnMetAspAsnLeu 20
      1      ATGAACTTCACAGTGGTTCACAGCCGCTGCTAGGGATGACACAGCATGACCAACTG 60
      21      GlnuGlnLeuLeuLeuGlyProLeuGluMetPheSerLysProValValIleLeu 40
      61      GAGAGGACGCTCATCTGCCCCATCTGCGTGGAGATGTTCTCCAAACAGTGGATCTTG 120
      41      ProGyGlnHisAsnLeuGlySerLysCysAlaAsnAspValPheGlnAlaSerAspPro 60
      121     CCGTGCACACACAACTGTCGCGCAATGTGCCACAGACGCTTCCAGGCTCGAATCT 180
      61      LeuTrpGlnSerArgLysSerThrValSerSerGlyValArgPheArgCysProSer 80
      181     CTATGGCAGTCCCGGGGGCTCCACCACTGTGTCTTCAGAGGCGGTTCCGCTGCCATCG 240
      81      CysArgHisGlnValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100
      241     TGCAGGCGTGAAGTGTCTCGACAGACAGCGGTCTACGCGCTGCGCAAACTGCTTA 300
      101     ValGlnAsnIleIleAspIleTyrLysGlnGlnSerSerArgProLeuHisAlaVal 120
      301     GTGGAGAACATTATGACATTTACAGAGAGGATCATCCAGGCGCTGCTCCAAAGCT 360
      121     GlnGlnHisLeuMetCysGlnGlnHisGlnAspGlnLysIleAsnIleTyrCysLeuSer 140
      361     GAGCAGCACTCATGTCGAGAGGATGAAGAAGATCATATTTACTGCTGAGC 420
      141     CysGlnValProThrCysSerLeuCysLysValPheGlnAlaHisLysAspCysGlnVal 160
      421     TGTGAGGTGCCACCTGCTCTCTGCAAGGTCTTGGTCCCAAGAGACTGTGAGGTG 480
      161     AlaProLeuProThrIleTyrLysArgGlnLys----- 171
      481     GCCCCAGTGCACCATTTACAAACGCCAGAAAGACAGATCTCACTCTGTGCCACAG 540
      171     ----- 171

```

```

Db      541     CTGAGTGCAGTGGCAACAACAACACTTAAGTGCAGGCTTGATGCCGAGCTCAAGTAT 600
QY      172     -----serGlnLeuSerAspGlyIle 178
      601     CTTCCATCTTAAGCTTCGACAACTAAGATTATGATGTGAGTCAAGTCCAGAGGATC 660
QY      179     AlaMetLeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGlnVal 198
      661     GCGATGCTGTGTGGCAGGCAATGACCGCGTCCAGACGATGATCACAGATGAGAGGTG 720
QY      199     CysGlnThrIleGlnAspAsnSerArgArgGlnLysGlnLeuAsnGlnArgPheGln 218
      721     TGCAGACTATGAGAGCAATAGCCGAGGACAGACAGAGAGAGAGAGAGAGAGAGAG 780
QY      219     ThrLeuCysAlaValLeuGlnGluArgLysGlyGlnLeuGlnAlaLeuAlaArgGln 238
      781     AGCTGTGCGCAGTGTGTGAGAGAGCCAAAGGAGAGTGTGTCAGGCGCGTGGCGGAG 840
QY      239     GlnGlnGlnLysLeuGlnArgValArgGlyLeuIleArgGlnTyrGlyAspHisLeuGln 258
      841     CAAGAGAGAGAGAGTGCAGCGCGCTCCGCGCTCATCGTCAATGATGCGACACTGAG 900
QY      259     GlySerSerLysLeuValGlnSerAlaIleGlnSerMetGlnGlnProGlnMetAlaLeu 278
      901     GCTCTCTTAAGCTGTGTGAGTCTGCCATCCATGAGAGAGAGAGAGAGAGAGAGAG 960
QY      279     TyrLeuGlnGlnAlaLysGlnLeuIleAsnLysValGlnValMetSerLysValGlnLeu 298
      961     TATCTCCACAGCGCCAAAGAGGCTGATCAATAGGTCGGGCGCATGTGCAGAGAGTGTG 1020
QY      299     AlaGlyArgProGlnProGlyTyrGlnSerMetGlnGlnPheSerValSerValGlnHis 318
      1021     GCAGGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
QY      319     ValAlaGlnMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGlnLysP 338
      1081     GTGGCCGAATGCTGTGGAGCAGCATCTTCAGGAGCGGCTTCCGGGAGAGAGAG-- 1137
QY      339     AspAspMetAlaLeuAspGlyGlnGlnGlyAsnAlaGlyLeuGlnGlnArgLeuAsp 358
      1138     ---GAGGTGCGCCCAAGAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1194
QY      359     ValProGlnGly 362
      1195     GGGCTTAAGGC 1206
      RESULT 4
      LOCUS      AX060632      1500 bp      DNA      linear      PAT 22-JAN-2001
      DEFINITION Sequence 54 from Patent WO0078954.
      ACCESSION      AX060632
      VERSION      AX060632.1 GI:12406053
      KEYWORDS
      SOURCE      human.
      ORGANISM      Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
      REFERENCE      1 (bases 1 to 1500)
      AUTHORS      Lal,P., Yue,H., Tang,Y.T., Baughn,M.R., Azimzal,Y. and Tran,B.
      TITLE      Human transcriptional regulator proteins
      JOURNAL      Patent: WO 0078954-A 54 28-DEC-2000;
      FEATURES
      source      1..1500
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
BASE COUNT      401 a      378 c      439 g      282 t
ORIGIN
Alignment Scores:
Pred. No.:      1.84e-93      Length:      1500
Score:          1261.50      Matches:      251
Percent Similarity: 89.04%      Conservative: 17

```



Best Local Similarity: 83.39% Mismatches: 27  
 Query Match: 65.98% Indels: 6  
 DB: 6 Gaps: 4

US-09-908-988b-2 (1-366) x AX060632 (1-1500)

QY 1 MetaspherhvalGlypPhelysProleuleuGlyAspAlaHisAsnMetaspasleu 20  
 DB 291 ATGAACCTCACAGTGGTTCCTCAAGCCGGCTAGGGGATGACACAGCATGACACACTG 350  
 QY 21 GlutylsGlnleuIleCysProIleCysleuGlnMetPheSerLysProValIleleu 40  
 DB 351 GAGAAAGCAGCTCATCTGCCCCCATCTGCGGAGATGTTCTCCAAACAGAGTGATCTTG 410  
 QY 41 ProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
 DB 411 CCGGCCAACACAACTGTGCGCGCAATGTGCCAAGAGCTCTCCAGGCGCTCGAATCCT 470  
 QY 61 LeutrgInserArgLysSerThrValSerSerGlyLysArgPheArgCysProSer 80  
 DB 471 CTATGGCAGTCCCGGGGCTCCACCACTGTCTTCAGAGAGCGGTTCCGCTGCCATCG 530  
 QY 81 CysArgHisGlnValValLeuAspArgHisGlnValTyrGlyLeuGlnArgAsnLeu 100  
 DB 531 TGCAGGCAATGAGTGTCTCTGGACAGACAGCGTGTCTACGCGCTGCAGCGAAACCTGCTA 590  
 QY 101 ValGlnAsnIleIleAspIleTyrLysGlnLysSerSerArgProleuHisAlaLysAla 120  
 DB 591 GTGGGAAACATTAATGACATTTACAAAGCGAGTCTCCAGCGCGCTGCACCTCCAGGCT 650  
 QY 121 GluGlnHisLeuMetCysGlnGlnHisGlnAspGlnLysIleAsnIleTyrCysLeuSer 140  
 DB 651 GAGCAGCACCTCATGTGGGAGAGCATGAAGAAGATCAATATTACTCCCTGAGC 710  
 QY 141 CysGlnValProThrCysSerLeuCysLysValPheGlyAlaHisLysAspCysGlnVal 160  
 DB 711 TGTGAGGTGCCACCTGCTCTCTGCAAGGTCTTCTGCGCCACAAAGGACTGTGAGGTG 770  
 QY 161 AlaProLeuProThrIleTyrLysArgGlnLysSerGlnLeuSerAspGlyIleAlaMet 180  
 DB 771 GCCCAGACGTCCACCATTTACAAAGCGCGAGAGTGTGAGCTCGCATGCGCATG 830  
 QY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGlnValCysGln 200  
 DB 831 CTGCTGGCAGGCAATGACCGCGCTGCAAGCATGATCACAAGATGAGAGGTGTGCAG 890  
 QY 201 ThrIleGlnAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGlnThrLeu 220  
 DB 891 ACTATCGAGGACAAATAGCCGAGGCGAGAAAGCATGTTTAAACCAAGAGTTTGAAGCGCTG 950  
 QY 221 CysAlaValIleuGlnGlnArgLysGlyGlnLeuGlnAlaLeuAlaArgGlnGln 240  
 DB 951 TGCCGAGTGTGAGAGGAGCGCAAGCGTGTGAGCTGTGAGCGCTGCGCGCGGAAACAGCA 1010  
 QY 240 uGluLysLeuGlnArg---ValArgLysLeuIleArg---GlnTyrGlyAspHisLeuGln 258  
 DB 1011 GGACAAAGCTTCAACCGCATCCAGCGGAGCTCAATTCGCGTCAATATGGCGAAGCAACTT 1070  
 QY 258 uGlySerSerLysLeuValGlnSerAlaIle-----GlnSerMetGlnLubProGlnMe 276  
 DB 1071 GGGAAAGCCCTCCCTGCTCAAAAGCCTGTGGAGGCTGCTCAAAATGCCAACTTT 1130  
 QY 276 tAlaLeuTyrLeuGlnGlnAlaLysGlnLeuIleAsnLysValGlyAlaMetSerLysVa 296  
 DB 1131 CCACATTTGATATACACAGAGC---CCAAACGAATAATGTGTGCGGACACTTACTATAAT 1187  
 QY 296 1 296  
 DB 1188 G 1188  
 RESULT 5  
 AX418850  
 LOCUS AX418850 2590 bp DNA linear PART 18-JUN-2002

DEFINITION Sequence 3 from Patent WO0206318.  
 ACCESSION AX418850  
 VERSION AX418850.1 GI:21523714  
 KEYWORDS  
 SOURCE  
 ORGANISM house mouse.  
 Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS Olson, E.N. and Spencer, J.A.  
 TITLE Methods and compositions for stabilizing microtubules and  
 JOURNAL intermediate filaments in striated muscle cells  
 BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)  
 FEATURES  
 source location/Qualifiers  
 1.2590  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 80..1717  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD35456.1"  
 /db\_xref="GI:21523715"  
 /translation="MSTSLNKSFSKREQTMDNLEKOLICPICLEMFTKPVYIIPCQ  
 NLCKKASDIFQASNPFLPTKGTTVASGGRFCPSCHREYVLDHRYGAGQRLVE  
 NIIDYRQESTRPBRKLDQPMCEHEERINITYCLNCEVPISLCIKVGARHDCQVAP  
 LTHVFOROKSELSIDGIALVGSNDNRVGVISQLEDTCTIECCRKQODLEKEDHL  
 YGILEERKTEMTQAITRQEEKLEHRLIRKSDHELNKSLVSGISGFQMDPEMAV  
 FIONAKTLILKTIKTVASAKFQMEKLEQGEINFTVYINREKIIIREIDFSREBE  
 DAGTIDEGREBDVVEEAEENVQIASGSEESLEKAEPSQLPRELDVAPPELDASS  
 PEPSPMPADVLVITQEVVIVIGSQQTQSEISQPSAEATADLPFTFSWKQGRKT  
 SSNPCKHSGSLQIQLIGLIGEDSSVQSAEVAEATNQAAVSGKSSSTAAATQIGF  
 EAPSPQOSALGSGGVILISQLATSSPSPLINSINE"  
 BASE COUNT 703 a 627 c 625 g 635 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,47e-85 Length: 2590  
 Score: 1162.00 Matches: 226  
 Percent Similarity: 76.99% Conservative: 55  
 Best Local Similarity: 61.92% Mismatches: 76  
 Query Match: 60.77% Indels: 8  
 DB: Gaps: 4  
 US-09-908-988b-2 (1-366) x AX418850 (1-2590)  
 QY 1 MetaspherhvalGlypPhelysProleuleuGlyAspAlaHisAsnMetaspasleu 20  
 DB 80 ATGACCACTTCTCTGAATTAACAAGTCTTCTCCAAAGACAGACCATGGATTAACCTTG 139  
 QY 21 GlutylsGlnleuIleCysProIleCysleuGlnMetPheSerLysProValIleleu 40  
 DB 140 GAAAGCAACATGATGTGTCGCCATGCTGCTAGAGATGTTTACAGAGCGCTGTGCTATCTG 199  
 QY 41 ProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
 DB 200 CCTTGCCAGCAACAACTGTGCGAGAAATGTGCCAGATCATCTTCCAGCGCTCTAACCCG 259  
 QY 61 LeutrgInserArgLysSerThrValSerSerGlyLysArgPheArgCysProSer 80  
 DB 260 TACTTACCACCAAGAGAGGAGCGACCGCTGCGATCAGGGGGCGGCTTCGCTGCTCCCTC 319  
 QY 81 CysArgHisGlnValValLeuAspArgHisGlnValTyrGlyLeuGlnArgAsnLeu 100  
 DB 320 TGCACACATGAGCTGTGTACAGACATGGGGTGTATAGCATGCGAGAGAACTGCTC 379  
 QY 101 ValGlnAsnIleIleAspIleTyrLysGlnLysSerSerArgProleuHisAlaLysAla 120  
 DB 380 GTGGAACAAATATGATATCTACAAAGCAGAAATCCACAGGCGCA---GAAAAAAATTTG 436  
 QY 121 GluGlnHisLeuMetCysGlnGlnHisGlnAspGlnLysIleAsnIleTyrCysLeuSer 140  
 DB 437 GACCAAGCCC---ATGTGTGAAGAGCATGAAGAGAAAGCATGATCAATCTATTGTGTGAAC 493

OY 141 CysGluValProThrCysSerLeuGlyValPheGlyAlaHisLysAspGluVal 160  
 |||||  
 Db 494 TGCAAGTGGCCACTGTTCTTGTGCAAGGTTTGGCCCATAGAGACTCCAGGTG 553  
 OY 161 AlaProLeuProThrIleTyrLysArgGlnLysSerGluLeuSerAspGlyIleAlaMet 180  
 |||||  
 Db 554 GCTCCCTGACATCATGTGTTCAGAGGAGCAAGTCAAGACTCAGTATGGTATGCTGTA 613  
 OY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetCgluValCysGln 200  
 |||||  
 Db 614 CTGTGGGAACCAAGATAGATGATCCAGGCTGATCAGCCACTGAGAGCACCTGTAAA 673  
 OY 201 ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnIleArgPheGluThr 220  
 |||||  
 Db 674 ACATTGAGGAGTCTCTGCAAGAAGCAAGCAAGCACTGTGAGAAATTTGATCACCTTA 733  
 OY 221 CysAlaValLeuGlnGluValArgLysGlyLeuLeuGlnAlaLeuAlaArgGlnGlu 240  
 |||||  
 Db 734 TAGCGCATCTCTGAGAGAGAGAGACTGAATGACCCAGCATCTGCAACACAGAGAG 793  
 OY 241 GluLysLeuGlnArgValArgGlyLeuIleArgGlnIleTyrGlyAspHisLeuGlySer 260  
 |||||  
 Db 794 GAGAACTGGAAACATGTCGAACTCTTATCAGAAATATTCGATCCTCGAGAACGTA 853  
 OY 261 SerLysLeuValGlnSerAlaIleGlnSerMetGluGluProGlnMetAlaLeuTyrLeu 280  
 |||||  
 Db 854 TCAGAGTTGGTGGAGTCAAGATCCAGTTCAATGATGAGCCCGAATGGCAGATTTCTG 913  
 OY 281 GlnGlnAlaLysGluLeuIleAsnLysValGlyAlaMetSerLysValGluLeuAlaGly 300  
 |||||  
 Db 914 CAGAAATGCCAAGACCCCTGTCACAAAGATCGTGAAGATCGTTCAGATGAGAG 973  
 OY 301 ArgProGluProGlyTyrGlySerMetGluGlnPheSerValSerValGluHisValAla 320  
 |||||  
 Db 974 AAACCTAGAACAAAGTTTATGATCATGACGCAACTTCACGTAATCTCAATGAGAGAA 1033  
 OY 321 GluMetLeuArgThrIleAspPheGlnProGlyAlaAlaGlyAspGluLysAspAsp 340  
 |||||  
 Db 1034 AAATATTCGTCGTAATTTGACTTT-----TCTAGAGAAAGAGAGAGAGAAAGAT 1084  
 OY 341 MetAlaLeuAspGlyGluGluGlyAsnAlaGlyLeuGluGluLysValPro 360  
 |||||  
 Db 1085 GCAAGAGAAATAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1135  
 OY 361 GluGlySerGlyLeu 365  
 |||||  
 Db 1136 GAGGCAAGAAATGTT 1150  
 RESULT 6  
 BC007750 1750 bp mRNA linear PRI 12-JUL-2001  
 LOCUS Homo sapiens, clone MGC:12836 IMAGE:4110783, mRNA, complete cds.  
 DEFINITION BC007750  
 ACCESSION BC007750.1 GI:14043531  
 VERSION 1  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 1750)  
 Strausberg, R.  
 Direct Submission  
 Submitted (11-MAY-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ARCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
 REMARK COMMENT  
 REMARK COMMENT

DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@nhgri.nih.gov](mailto:nisc_mgc@nhgri.nih.gov)  
 Shevchenko, Y., Wechterby, K.D., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Balesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McLooney, J.C.,  
 McDowell, J., Pearson, R., Snyder, B., Stantrop, S., Thomas, P.J.,  
 Tjongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,  
 Zhang, L.-H. and Green, E.D.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>  
 Series: IRAL Plate: 17 Row: d Column: 4  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis, Similarity but not identity to protein.  
 Location/Qualifiers  
 1..1750  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:12836 IMAGE:4110783"  
 /tissue\_type="Muscle, rhabdomyosarcoma"  
 /clone\_id="NIH-MGC\_17"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"  
 190..1500  
 /codon\_start=1  
 /product="Unknown (protein for MGC:12836)"  
 /protein\_id="AAH07750.1"  
 /db\_xref="GI:14043532"  
 /translation="MDNEKOLICPLEMTFKPVYILPCQNLCKRCASDIQASNP  
 ILPTRGTTMASGGRFCPSCHREYVLDHGVYGLDRLNLVNIIDLYKESIRPERK  
 SDPWCSEHEERINITYLCNCPYSCSLKVGANHCDOYAPLTHFORKSEISDGI  
 AILGVNDKVOGVISQLEDCTIECCCKROKOELEKEDYGLILEERNEMTOVT  
 RTQEKLEHVALIRKYSDBHLENSKLVSGIOFMEPEMAVFLONAKTLTKISGS  
 KAFQWEKLEHGVENNHTVNLNREKIIREDIDYREDEDEEEREGEGEKREKPVG  
 GEAYEVEVNVQTEFRBEDENPERKASELSQVYLOAPALPAPSSSEPPALPPADA  
 PVTQIGFPAFPDGGAAAPASGSADSEPARHIFSPWLSLNE"  
 ORIGIN  
 BASE COUNT 556 a 389 g 433 t 372 c  
 Alignment Scores:  
 Pred. No.: 3.28e-85 Length: 1750  
 Score: 1161.00 Matches: 221  
 Percent Similarity: 75.48% Conservative: 56  
 Best Local Similarity: 60.22% Mismatches: 74  
 Query Match: 60.72% Indels: 16  
 DB: Gaps: 3

US-09-908-988b-2 (1-366) x BC007750 (1-1750)  
 OY 1 MetAspPheThrValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20  
 |||||  
 Db 142 ATGAGCGCATCTCTGAAATTAACAATCTTTTCAAGAGAGAGAGAGAGAGAGAGAGATTA 201  
 OY 21 GluLysGlnLeuIleLecysProIleCysLeuGlnMerPheSerLysProValIleLeu 40  
 |||||  
 Db 202 GAGAGCAACCTCATCTGTCCTTGAAGATGTTTCAAGAACTGTGGGATTTCTC 261  
 OY 41 ProCysGlnHisAsnLeuGlySerGlyCysAlaAsnAspValPheGlnAlaSerAspPro 60  
 |||||  
 Db 262 CTTGTGCGCACACACTGTGTGGAATGTGCCAGATGATTTTCCAGGCTCTTAACCG 321  
 OY 61 LeuTyrGlnSerArgGlySerThrThrValIleSerGlyGlyArgPheArgCysProSer 80  
 |||||  
 Db 322 TATTTGCCACACAAGAGAGGTACACACATGATGATGAGGGGCGGATTCGCCCATTC 381  
 OY 81 CysArgHisGluValIleLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100  
 |||||  
 Db 382 TGTAGACATGAAGTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 441





```

Db 433 GTGGAATATCATGATGATACATGACAGAGAGGATCCACAGAGCA-----GAAAGAAA 486
Oy 121 GUGUHHSLMEUCYSGUGUHHSLGUSPGLUYSLEANSILETYRYSLEUSER 140
Db 487 TCCGACGACCCATGCTGGAGAGACATGAAAGAGAGCCGATCAACTACTCTCGAAC 546
Oy 141 CysgluValProthCysSerLeuCySLysValPheGluAlaHisLysAspCysgluVal 160
Db 547 TCCGAGATGACCCACTGCTCTCTCTGTCAGAGGTGTTGGTGCACACAAAGACTCCAGGTG 606
Oy 161 AlaProLeuProthIleTyLysArgGlnLysSerGluLeuSerAspGlyIleAlaMet 180
Db 607 GCTCCCTCACACATGCTGTTCAGAGACAGAGTGTAGCTGAGTGTGTCATGCCATCGCATC 666
Oy 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThGlnMetGluGluValCysGln 200
Db 667 CTCTGGGCGACGACGATGCTGCTCCGAGGAGTGCATGACGCTGGAAGACACCTCGCAA 726
Oy 201 ThrIleGluAspAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGluThrLeu 220
Db 727 ACTATCGAGGATGTTGCGAAGAAACAGAAACAGAGCTTTGTGAGAGTTTGATTACCTG 786
Oy 221 CysAlaValLeuGluGluValArgLysGlnLeuLeuGlnAlaLeuAlaArgGluGln 240
Db 787 TATGCAATTTTGGAGAGAGAGAAATGAAATGACCCAGTCAATCCGACCAAGAGAG 846
Oy 241 GluLysLeuGlnArgValArgGlyLeuIleArgGlnTyGlyLysPheGluGlySer 260
Db 847 GAGAAACCTGGAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906
Oy 261 SerLysLeuValGlnSerAlaIleGlnSerMetGlnGlnProGlnMetAlaLeuTyL 280
Db 907 TCAAGCTTGTTGAGTACGAGCAATTCAGTTATGATGATGACCAAAATGCGAGTCTCTG 966
Oy 281 GlnGlnAlaLysGlnLeuIleAsnLysValGlyAlaMetSerLysValGlnLeuAlaGly 300
Db 967 CAGATGCGCAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1026
Oy 301 ArgProGlnProGlyTyLysLeuSerMetGlnGlnPheSerValSerValGlnHisValAla 320
Db 1027 AAAAATGAAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1086
Oy 321 GlnMetLeuArgThrIleAspPhe-----GlnProGly 331
Db 1087 AAGATTAATACGTAATTTACTTTTACAGAGAAATGAAATGAAATGAAATGAAATGAAATG 1146
Oy 332 AlaAlaGlyAspGlnLysAspAspMetAlaLeuAspGlyGlnGluGlyAsnAlaGly 351
Db 1147 GGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1191
Oy 352 LeuGluGluGluGluArgLeuAsp 358
Db 1192 GTAGAAGTGAAGAGAGTGAAG 1212

RESULT 9
LOCUS AK091310 2634 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ33991 fis, clone DFNES2007332, moderately
similar to Mus musculus RING-finger protein MURF mRNA.
ACCESSION AK091310
VERSION AK091310.1 GI:21749650
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens normal dermal fibroblasts (Neonatal Skin) (NHDF2564)
CDNA to mRNA, clone:lib:DFNES2 clone:DFNES2007332.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE
1..2634
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DFNES2007332"
/cell_type="normal dermal fibroblasts (Neonatal Skin)"
(NHDF2564)
/clone_lib="DFNES2"
/notes="Cloning vector: pME18SFL3-primary culture, normal
dermal fibroblasts"
BASE COUNT 793 a 580 c 675 t
ORIGIN
Alignment Scores:
Pred. No.: 5.5e-85 Length: 2634
Score: 1161.00 Matches: 221
Percent Similarity: 75.48% Conservative: 56
Best Local Similarity: 60.22% Mismatches: 74
Query Match: 60.72% Indels: 16
DB: Gaps: 3
US-09-908-988b-2 (1-366) x AK091310 (1-2634)
Oy 1 MetAsnPheThrValGlyPheLysProLeuLeuGlyAspAlaHisAsnMetAspAsnLeu 20
Db 377 ATGAGCGCATCTCTGAATTAACAATCTTTTCCAAAGACAGACAGACATGATTAACCTTA 436
Oy 21 GluLysGlnLeuIleCysProIleCysLeuGlnMetPheSerLysProValValIleLeu 40
Db 437 GAGAGCAACTCATCTGTCGCCATCTGCTTGAAGATGTTCAAGAAACCTGCGTGAATCTC 496
Oy 41 ProCysGlnHisAsnLeuCysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60
Db 497 CTTTCTCAGCAACACCTGTGTGGAATGTGCCAGAGATTTTCCAGGCTTCAACCCG 556
Oy 61 LeuTrpGlnSerArgGlySerThrThrValSerSerGlyIleArgPheArgCysProSer 80
Db 557 TATTGCGCCACAGAGAGAGTACCCATGCGATCGAGGGGGCGGATTCGCTCCCATCC 616
Oy 81 CysArgHisGluValValLeuAspArgHisGlyValTyGlyLeuGlnArgAsnLeuLeu 100
Db 617 TGTAGACATGAGAGTGTGTTGATAGACATGGGGGTATGAGACTTCAGAGAACCTGCTG 676
Oy 101 ValGluAsnIleIleAspIleTyLysGlnGlnLysSerArgProLeuHisAlaLysAla 120
Db 677 GTGGAATATCATGATGATACATGACAGAGAGTCCACAGAGCA-----GAAAGAAA 730
Oy 121 GUGUHHSLMEUCYSGUGUHHSLGUSPGLUYSLEANSILETYRYSLEUSER 140
Db 731 TCCGACGACCCATGCTGGAGAGACATGAAAGAGAGCCGATCAACTACTCTCGAAC 790
Oy 141 CysgluValProthCysSerLeuCySLysValPheGluAlaHisLysAspCysgluVal 160

```

```

Db 791 TCGCAAGTACCACCTGCTCTCTGTGCAAGGTGTTGGTGACACAAAGACTGCCAGGTG 850
Qy 161 AlaProleupProthrlleTyrlsArgGlnLysSerGlnLeuSeraspGlyIleAlaMet 180
Db 851 GCTCCCTCAGCAGTGTGTCCAGACAGCAAGAACTGTGAGCTCATGTGATGGCATGGCCATC 910
Qy 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGlnGluValCysGln 200
Db 911 CTCGGGGCAGCAGACATCGATCGATCCAGGAGTGTATCAGCCAGCTGGAAGACACTGCCAAA 970
Qy 201 ThrIleGlnAspAsnSerArgArgGlnLysGlnLeuLeuasnGlnArgPheGlnThrLeu 220
Db 971 ACTATCGAGGAGTGTGCGAGAAACAGAAACAGACTTGTGTGAAGATTGATTTACTCTG 1030
Qy 221 CysAlaValLeuGlnGluValArgLysGlnLeuLeuGlnAlaLeuAlaArgGlnGln 240
Db 1031 TATGGCATTTGGAGAGAGAGAAATGAATGACCAACTCATTTACCCGACCCCAAGAG 1090
Qy 241 GlnLysLeuGlnArgValArgGlyLeuIleArgGlnTyrlsPheIleGlnGlySer 260
Db 1091 GAGAAACTGGAACATGTCTGCTCTGTCAAAAAGTATCTGATCATTTGGAGAACGTC 1150
Qy 261 SerLysLeuValGlnSerAlaIleGlnSerMetGlnGlnProGlnMetAlaLeuTyrlen 280
Db 1151 TCAAAAGTTGGTTGAGTCAAGAAATTCAGTTATGATGATGAGCCAGAAATGCGAGTGTCTG 1210
Qy 281 GlnGlnAlaLysGlnLeuIleAsnLysValGlnAlaMetSerLysValGlnLeuAlaGly 300
Db 1211 CAGATGCGCAAAACCTCGTAAAAAATCTCGGAGACATCAAAAGCATTTTCAGATGGAG 1270
Qy 301 ArgProGlnProGlyTyrlsSerMetGlnGlnPheSerValSerValGlnHisValAla 320
Db 1271 AAAATAGAACATGCTATGATGAGAACATGACCACTTCAACCTCAATAGAGAGAA 1330
Qy 321 GlnMetLeuArgThrIleAspPhe-----GlnProGly 331
Db 1331 AAGATTAATACGTGAATATGACTTTTACAGAGAAATGATGAACATGAAAGAGAAAGGC 1390
Qy 332 AlaAlaGlyAspGlnGlnAspAspMetAlaLeuAspGlyGlnGlnGlyAsnAlaGly 351
Db 1391 GGAGAGAGAGAAAAGAAAGGA-----GAAGAGAAAGTGGAGAGAGAGAA 1455
Qy 352 LeuGlnGlnGlnArgLeuAsp 358
Db 1436 GTAGAAAGTGAAGAGGTAGAA 1456

RESULT 10
LOCUS HSA243488 1810 bp mRNA linear PRI 29-JUN-2001
DEFINITION Homo sapiens titin zinc-finger anchoring protein, 50kDa isoform.
ACCESSION AJ243488
VERSION AJ243488.1 GI:14588845
KEYWORDS alternative splicing; signal transduction; titin zinc-finger anchoring protein; titian.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1810)
AUTHORS Kelly,R., Neubauer,G. and Gautel,M.
TITLE A novel RING finger protein associated with titin kinase
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1810)
AUTHORS Gautel,M.S.
TITLE Direct Submission
JOURNAL Submitted (29-JUN-1999) Gautel M.S., Physikalische Biochemie, Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse 11, Dortmund, 44227, GERMANY
FEATURES
source
1..1810
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="cardiomyocyte"

```

```

CDS
227..1585
/feature="cardiac muscle"
/function="putative role in signal transduction"
/size="50kDa isoform"
/codon_start=1
/evidence=experimental
/product="titin zinc-finger anchoring protein"
/db_xref="GI:14588846"
/translation="MSASLNKSFKSKQDMNLEKOLICPICLEMFTKRVVILPCOH
NICRKAASDFQASNFTLPTRGGTTMASGGRFCPSCHNEVVDHNGVGLQRLVIG
NIDYIKQESTPREKKSDDPCHEHEBERINICLNEVPLSCSKYFGARKDCQVAP
LHVFQKRSSELDGIALIVGSDNDROGVISQLEDCTKTECCRKQKOECEKEDYL
YGLIEERKNEMTOVITRTOEKLLEHVALILKTYSDHLENYKSLVYESGLOFDEPMAY
FLONAKTLTKISSEASKAFQMEKIEHNEEMNHNTVUNIREKTIITREIDFYRDEDEE
EEEGGGEKEEENGEENAEVFEVENVQEFREDENPEKASELSOVELDAAREALP
VSSPEPPRLPPAADAPVTDIGFAPLQDQAAAPASGSGADSEPARHIFSFWSLNL
NE"

BASE COUNT 557 a 420 c 448 g 385 t
ORIGIN

Alignment Scores:
Pred. No.: 7,23e-85 Length: 1810
Score: 1157.00 Matches: 222
Percent Similarity: 76.60% Conservative: 53
Best Local Similarity: 61.84% Mismatches: 74
Query Match: 60.51% Indels: 10
DB: 9 Gaps: 3

US-09-908-988b-2 (1-366) x HSA243488 (1-1810)
Qy 1 MetAspPheThrValGlnPheLysProLeuLeuGlnLysPalaHisAsnMetAspAsnLeu 20
Db 227 ATGAGCGCATCTGTGAATTCAAATCTTTTCCAAAGACGACGACACCATGATCACTTA 286
Qy 21 GlnLysGlnLeuLeuLecGlyProLecGlyLeuGlnLecMetPheSerLysProValAlaLeu 40
Db 287 GAGAACCACTCATCTGTCCATCTGCTTAGAGATGTTACAGAAACCTGTGATCTCTC 346
Qy 41 ProCysGlnHisAsnLecCysArgLysCysAlaLysAspValPheGlnAlaSerAsnPro 60
Db 347 CCTTGTACACCACTGTGTAGGAAATGTGCCAGTATATTTTCCAGGCCCTGTAAACCCG 406
Qy 61 LeuTrpGlnSerArgLysSerThrThrValSerSerGlyGlyArgPheArgCysProSer 80
Db 407 TATTTGCCCAAGAGAGAGTACCACCATGCGATCAGGGGCCCATTTCCGTGCCCATCC 466
Qy 81 CysArgHisGlnValValLeuAspArgHisGlyValTyrlsLeuGlnAlaArgAsnLeu 100
Db 467 TGTAGACATGAAGTGGTTTGGATAGACATGGGGTATATGATGATCAGAGAACCTGCTG 526
Qy 101 ValGlnAsnIleLeuAspIleTyrlsGlnGlnSerSerArgProLeuHisAlaLysAla 120
Db 527 GTGGGAATATCATTTTACATCTTCAACGACGAGTCCACCGGCA-----GAAAAGAAA 580
Qy 121 GlnGlnHisLeuMetCysGlnGlnHisGlnAspGlnLysIleAsnIleTyrlsCysLeuSer 140
Db 581 TCCGACACCCCATGTGCGAGAGAAATGAAGAGAGGCGCATCACTACTGTGTGAC 640
Qy 141 CysGlnValProThrCysSerLecCysLysValPheGlnAlaHisLysAspCysGlnVal 160
Db 641 TCGGAAGTACCACCTGCTCTGTGCAAGGTGTTGGTGACACAAAGACTGCCAGGTG 700
Qy 161 AlaProleupProthrlleTyrlsArgGlnLysSerGlnLeuSerArgPlyIleAlaMet 180
Db 701 GCTCCCTCAGCAGTGTGTCCAGACAGCAAGAACTGTGAGCTCATGTGATGGCATGGCCATC 760
Qy 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGlnGluValCysGln 200
Db 761 CTCGGGGCAGCAGACATCGATCGATCCAGGAGTGTATCAGCCAGCTGGAAGAGAACTGCCAAA 820
Qy 201 ThrIleGlnAspAsnSerArgArgGlnLysGlnLeuLeuasnGlnArgPheGlnThrLeu 220

```

Db 821 ACTATGAGCAATGTTGCAAGAAAACAGAAAACAGAGCTTTGTGAGAAATTGATTACCTG 880  
OY 221 CysAlaValLeuGluGluArgLysGlyGluLeuLeuGlnAlaLeuAlaArgGluGln 240  
Db 881 TATGCAATTTTGGAGAGAGAGAAATGAAATGACCAGCAATTCATCCACCAACCAAGAG 940  
OY 241 GluLysLeuGlnArgValArgGlyLeuLeuLeuArgGlyLeuGlySer 260  
Db 941 GAGAAACTGGAACATGTCCTGCTGTGATCAAAAAGTATTCATCTGTTGGAGAACGTC 1000  
OY 261 SerLysLeuValGluSerAlaLeuGlnSerMetGluGluProGlnMetAlaLeuTyrLeu 280  
Db 1001 TCAAGTTGTTGAGTACGAGCAATTCAGTTATGATGAGCAGCAAAATGCGAGTCTTCTG 1060  
OY 281 GlnGlnAlaValGluLeuLeuLeuLysValGlyAlaMetSerLysValGluLeuAlaGly 300  
Db 1061 CAGAAATGCCAAAACCTGCTGTAATAAAATCTCGAAGCAATCAAGCAATTCAGATGAG 1120  
OY 301 ArgProGluProGlyTyrGluSerMetGluGlnPheSerValSerValGluHisValAla 320  
Db 1121 AAAATAGAACATGCGTATGAGAACATGACCACTTCACACTCACTCATAGAGAA 1180  
OY 321 GluMetLeuValArgThrLeuAspPheGlnProGlyAlaAlaGlyAspGluGluAspAsp 340  
Db 1181 AAGATAATACGTGAATATGACTTTTAC-----AGAGAAATGATGAGAT 1222  
OY 341 MetAlaLeuAspGlyGluGluGluAsnAlaGlyLeuGluGluGluGluGluGluGlu 359  
Db 1223 GAA-----GAAAG 1273  
RESULT 11  
LOCUS HSA243489 2098 bp mRNA linear PRI 29-JUN-2001  
DEFINITION Homo sapiens titin zinc-finger anchoring protein, 60kDa isoform.  
ACCESSION AJ243489.1 GI:14588847  
VERSION 1  
KEYWORDS alternative splicing; signal transduction; titin zinc-finger anchoring protein; fizian.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 2098)  
Kellly,R., Neubauer,G. and Gautel,M.  
TITLE A novel RING finger protein associated with titin kinase  
JOURNAL Unpublished  
AUTHORS Gautel,M.S.  
TITLE Direct Submision  
JOURNAL Submitted (29-JUN-1999) Gautel M.S., Physikalische Biochemie,  
Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse  
11, Dortmund, 44227, GERMANY  
FEATURES  
source  
1..2098  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_type="cardiomyocyte"  
/tissue\_type="cardiac muscle"  
227..1873  
/function="putative role in signal transduction"  
/note="60kDa isoform"  
/evidence="experimental"  
/product="titin zinc-finger anchoring protein"  
/protein\_id="CAC43020.1"  
/db\_xref="GI:14588848"  
/translation="MSASLNYSFSKEQDTMDNLEKOLICPICLEMFTKPVLLPCOH  
NLCRCASDIPQASNPYLPRTGRTMASGGRFCRPHVVDLRHGVYGLORLLVG  
NIDYIKOSTREPKSDOPMCEHEBERINIVLCEVETGSCIKFVGHAKOCQVAP  
LTHVROKSELSGIALIVSGNDROGVISOLEPDKTEPCCRKOKQKCEPFDL  
YGLIERKEMTOVITRTOEKLEHVALIKTSIDLENSKLVESSTIQWDEPEMAY  
FLQNAKTLTKRISEASKAFQMEKIEHGYENMHFTVNLNREKTIIRIDYRDEDEE  
EEEGEGEKEBEGEVEGAEVVEEVENVOTEFPEDEENPEKASELSQVELQAAFGALP

VSSPEPPALPPADAPVTOGEVVPPTGSEQTSETPYPAAEIADPLFPSPWKGOT  
RKATNPCTGSEGLIGIPGSEDSVVRKAEVAAAASERAAVSSEKTSAPATISO  
IGFPAIPLOGAAPAPSSGADSPRARHIFPSWLSNLE"  
BASE COUNT 632 a 497 c 534 g 435 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.72e-85 Length: 2098  
Score: 1157.00 Matches: 222  
Percent Similarity: 76.60% Conservative: 53  
Best Local Similarity: 61.84% Mismatches: 74  
Query Match: 60.51% Indels: 10  
Gaps: 3  
US-09-908-988b-2 (1-366) x HSA243489 (1-2098)  
OY 1 MetAsnPheThrValGlyPheLysProLeuGluGlyAspAlaHisAsnMetAspAsnLeu 20  
Db 227 ATGAGCGCATCTCTGAATTAACAATCTTTTCCAAAGACGACGACACCATGCAATCTTA 286  
OY 21 GluLysGlnLeuLeuLeuLysProLecysLeuGluMetPheSerLysProValAlaLeu 40  
Db 287 GAGAGCAACTCATCTGTCCTGCTTGAAGAGTTCACGAAACCTGTGTGATTCTC 346  
OY 41 ProCysGlnHisAsnLeuLysArgLysCysAlaAsnAspValPheGlnAlaSerAsnPro 60  
Db 347 CCTTTCAGACCAACACTGTGTGAGAAATGTCCACGATATTTTCCAGGCTTAACCCG 406  
OY 61 LeuTrpGlnSerArgLysSerThrThrValSerSerGlyLysArgPheArgCysProSer 80  
Db 407 TATTTGCCCAAGAGAGAGTACCCACCATGCAATCAGGGGGCGGATCCGCTCCCATCC 466  
OY 81 CysArgHisGluValValLeuAspArgHisGlyValTyrGlyLeuGlnArgAsnLeu 100  
Db 467 TGTAGACATGAGTGTGTTGATGATACATGGGGTATGAGACTTTCAGAGAACTGTG 526  
OY 101 ValGluAsnLeuLeuAspLecTyrLysGlnGluSerSerArgProLeuHisAlaLysAla 120  
Db 527 GTGGAAATATCATTTGACTTACATCAAGCAGAGTCCACCAAGCCA-----GAAAGGAA 580  
OY 121 GlnGlnHisLeuMetCysGluGluHisGluAspGluLysLysLysLecTyrCysLeuSer 140  
Db 581 TCCGACCAAGCCCATGTGCGCAGAGAACATGAAAGAGACGCAATCACTGCTCGAAC 640  
OY 141 CysGluValProThrCysSerLeuLysValPheGlyAlaHisLysAspCysGluVal 160  
Db 641 TCGAAGTACCCACCTGCTCTGTGCAAGGTGTTGGTCACACAAAGACTCCAGGTG 700  
OY 161 AlaProLeuProThrLecTyrLysArgGlnLysSerGluLeuSerAspGlyLeuAlaMet 180  
Db 701 GCTCCCTCACCACATGCTGTCCAGACAGAAAGTGTGAGCTCATGATGCAATCGCATC 760  
OY 181 LeuValAlaGlyAsnAspArgValGlnAlaValIleThrGlnMetGluValCysGln 200  
Db 761 CCGTGGCGAGCAACGATGAGTCCAGGAGTATCAGCCAGCTGGAAGACACCTCGCAA 820  
OY 201 ThrLecLysAsnSerArgArgGlnLysGlnLeuLeuAsnGlnArgPheGluThrLeu 220  
Db 821 ACTATGAGAGAGTGTGCAAGAAACAGAAACAGAGCTTTGTAGAGATTGATTACCTG 880  
OY 221 CysAlaValLeuGluGluArgLysGlyGluLeuGlnAlaLeuAlaArgGluGln 240  
Db 881 TATGCAATTTTGGAGAGAGAGAAATGAAATGACCAGATCATTCACCAACCAAGAG 940  
OY 241 GluLysLeuGlnArgValArgGlyLeuLeuLeuArgGlyLeuGlySer 260  
Db 941 GAGAAACTGGAACATGTCCTGCTGTGATCAAAAAGTATTCATCTGTTGGAGAACGTC 1000  
OY 261 SerLysLeuValGluSerAlaLeuGlnSerMetGluGluProGlnMetAlaLeuTyrLeu 280  
Db 1001 TCAAGTTGTTGAGTACGAGCAATTCAGTTATGATGAGCAGCAAAATGCGAGTCTTCTG 1060  
OY 281 GlnGlnAlaValGluLeuLeuLysValGlyAlaMetSerLysValGluLeuAlaGly 300



Db	1061	CAGATGCGAACCCTGTATAAAATTCGACAGCATCAAGCATTCAGATGGAG	1120
Qy	301	ArgProGluTyrGlyTyrGluSerMetGluInpHisSerValSerValGluHisValAla	320
Db	1121	AAAATGAACAATGCCTATGACGACATGAACACTTCACAGTCACACCTCATAATGAGAAGA	1180
Qy	321	GluMetLeuArgThrIleAspPheGlnProGlyAlaIleAlaGlyAspGluAspAsp	340
Db	1181	AAGATAAATGCTGAATTGACTTTTAC-----AGAGAACATGACAGAT	1222
Qy	341	MetaLeuAspGlyGluGluGlyYasnalaGlyLeuGluGluArgLeuAspVal	359
Db	1223	GAA-----GAAGAAGAAGAACGCCGAGAACAGCAAAAAAGAAAGAAAGAGAGAGTG	1273
RESULT 12			
LOCUS	BC015717	1211 bp	mRNA linear PRI 11-OCT-2001
DEFINITION	Homo sapiens, similar to ring finger protein 28, clone MGC:17320		
ACCESSION	IMAGE:3922363, mRNA, complete cds.		
VERSION	BC015717.1		
KEYWORDS	BC015717.1 GI:16041696		
SOURCE	MGC.		
ORGANISM	Homo sapiens. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1211) Strausberg,R. Direct Submission Submitted (09-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REFERENCE	NIH-MGC Project URL: http://mgc.nci.nih.gov		
AUTHORS	Contact: MGC help desk		
TITLE	Email: gcaps-remail.nih.gov		
JOURNML	Tissue Procurement: ATCC/DCTP/DIP		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Genome Sequence Centre,		
	BC Cancer Agency, Vancouver, BC, Canada		
	Info@cgsc.bc.ca		
	Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Kirzysinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saedi, Jacqueline Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranda Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Maria.		
REMARK	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAX Plate: 15 Row: o Column: 19 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 14211914. Location/Qualifiers		
FEATURES	1..1211		
SOURCE	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="MGC:17320 IMAGE:3922363"		
	/tissue_type="Skin, melanotic melanoma."		
	/clone_id="NIH_MGC_72"		
	/lab_host="DH10B"		
	/note="Vector: PCMV-SPORT6"		
	135..1175		
	/codon_start=1		
	/product="Similar to ring finger protein 28"		
	/protein_id="AAH15717.1"		
	/db_xref="GI:16041697"		
	/translation="MDYKSSLIIDGNMENLEOLICPTCLEMFTRPVILPOHNIC RKANDIFQANPYWTSRGSVSMGSGRFPCPRCHEVIIMDRGVYGLONNLVENI		
CDS			

[illegible]







```
/note="putative alternative translation initiation codon"
/evidence=no!_experimental
475..1497
/gene="RNF28"
475..1497
/gene="RNF28"
/note="short isoform"
/codon_start=1
/product="ring finger protein 28"
/protein_id="CAC33173.1"
/db_xref="GI:13171051"
/translation="MENTEROLICPLEMPETKRVVILPCOHNCRKANIFQANP
YTRSRSSVSMSSGRPCPTCRHEVINDRNVITGLRNLVENTIDIKQCSSPLQ
KSHPKCKEHEDEKINYLCEVPTCSKVFGIHAKVAPQSVFGQKTELNC
ISMLVANDNRFVQIITOLEDSRYTKENSQKVELSOKEDTLAIIDERSKELORI
TOCEKERSIFEALIOQYOLEDKSTKLETAIVELDEPGGATFLITAKOLIKSIVA
SKGCOGKTQGFENMDFLDEHIDALRAIDFGDEDEEERIEEDQDEEESTES
KEEGHO"
```

BASE COUNT 565 a 460 c 604 g 468 t

ORIGIN

Alignment Scores:

Pred. No.:	2,33e-80	Length:	2097
Score:	1102.50	Matches:	211
Percent Similarity:	78.00%	Conservative:	62
Best Local Similarity:	60.29%	Mismatches:	66
Query Match:	57.66%	Indels:	11
DB:	9	Gaps:	4

US-09-908-988b-2 (1-366) x HSA291713 (1-2097)

QY 10 LeuLeuGLyAspAlaHisAsnMetAspAsnLeuGLyGlnLeuIleCysProIleCys 29

Db 454 CTGATCCAGATGGAGTCCATCCATGGAGAACTTGAGAGACGAGCTGATCGCCCTATCGC 513

QY 30 LeuGLMetPheSerLysProValValIleLeuProCysGlnHisAsnLeuCysArgLys 49

Db 514 CTGGAGATGTTTACCAAGCCAGTGGTCACTTGGCCGACGACACACCTGCGCGGAG 573

QY 50 CysAlaAsnAspValPheGlnAlaSerAsnProLeuTPGInSerArgLysSerThr 69

Db 574 TGTGCCAATACATCTTCCAGGCTGCAATCTCTACTGGACACGGCGGCGGCTGAGTG 633

QY 70 ValSerSerGlyLysArgPheArgCysProSerCysArgHisGluValValLeuAspArg 89

Db 634 TCCATGCTGGAGGCCCTTCCGCTGCCCTGCCCTGCCCTGCCCTGCCCTGCCCTGCT 693

QY 90 HisGLValLysArgLysLeuGlnArgAsnLeuValGlnAsnIleIleAspLysIleArgLys 109

Db 694 CAGGAGTGTACGCGCTGCAAGGACCTGCTGGTGGAGAACATCATCATCAACAA 753

QY 110 GlnGLu---SerSerArgProLeuHisAlaLysAlaGlnGlnHisLeuMetCysGlnGlu 128

Db 754 CAGAGAGTCTCCAGTGGCGCGCTG-----CAGAAAGGCACTCACCCCATGTGCAAGAG 807

QY 129 HisGLuAspLysLysIleAsnIleTyrCysLeuSerCysGlnValProThrCysSerLeu 148

Db 808 CACGAGATAGAGAAATCAACATCTACTGTCTACGTGTGAGTGCCACCTGCTCCATG 867

QY 149 CysLysValPheGlnAlaHisLysAspCysGlnValAlaAlaProLeuProThrIleArgLys 168

Db 868 TGCAGAGTGTGGGATCCACAAAGGCTGCGAGGTGGCCCATTTGCACAGTCTTCCAG 927

QY 169 ArgGLnLysSerGlnLeuSerAspGlyIleAlaMetLeuValAlaGlyAsnAspArgVal 188

Db 928 GGACAAAGATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 987

QY 189 GlnAlaValIleThrGlnMetGlnGlnValCysGlnThrIleGlnAspAsnSerArgArg 208

Db 988 CAGACCATCATCATCTGAGTGGAGATTCCTCCGAGTACCAAGAGAGACAGTACACAG 1047

QY 209 GlnLysGlnLeuAsnGlnArgPheGlnThrLeuCysAlaValLeuGlnGlnArgLys 228

Db 1048 GTAAGGAGAGCTGACGCCAGAACTTGTACAGCTTGTATGCCATCTGGATGAGAGAAA 1107

QY 229 GLYGLuLeuGlnAlaAlaAlaArgGlnGlnGlnGlnLysLeuGlnArgValArgLys 248

Db 1108 AGTGAGTGTCTGACGCGGATACCGACGAGAGAGGAGAGAAAGCTTACCTTCATCGAGGCC 1167

QY 249 LeuIleArgGlnTyrGlyAspHisLeuGlnGlnGlnGlnLysSerSerLysValAlaGlnSerAlaIle 268

Db 1168 CTGATCCAGAGATACAG 1227

QY 269 GlnSerMetLysGlnProGlnMetAlaLeuTyrLeuGlnGlnAlaLysGlnLeuIleAsn 288

Db 1228 CAGTCTCCCTGAGCAGCCTGGGGGAGCCACCTTCCCTTGAATGCCAAGCACTCATCAA 1287

QY 289 LysValGlyAlaMetSerLysValGlnLeuAlaGlyLysArgProGlnProGlyTyrGlnSer 308

Db 1288 AGCATTTGAGAGCTTCCAGAGGCTGCCAGCTGGGAGAGACAGAGAGAGAGAGAGAGAG 1347

QY 309 MetGlnGlnPheSerValSerValGlnHisValAlaGlnMetLeuArgThrIleAspPhe 328

Db 1348 ATGACATCTTCTTACTTGTGATTTAGACACATACAGAGCGCCCTGAGAGCCATTCATCTT 1407

QY 329 GlnProGlyAlaAlaGlyAspGlnGlnLysAspAspMetAla-----LeuAsp 344

Db 1408 -----GGACAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1455

QY 345 GLYGLuGlnLysAlaGlyLysLeuGlnGln 354

Db 1456 GAGGAGAGATCCACAG 1485

RESULT 16

HSAG276484 1756 bp mRNA linear PRI 04-JAN-2002

LOCUS HSAG276484

DEFINITION Human mRNA for muscle specific RING finger 2 (MURF2) protein (MURF2 gene).

ACCESSION AJ276484 GI:18073355

VERSION AJ276484.1

KEYWORDS MURF2 protein; RING finger protein.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1756)

AUTHORS Stanchi, F.

TITLE Characterisation of MURF2, a new muscle-specific RING finger protein of the RbC family that associates with microtubules

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1756)

AUTHORS Stanchi, F.

TITLE Direct Submission

JOURNAL Submitted (20-MAR-2000) Stanchi F., CRIBI Biotechnology Centre, Universita di Padova, Via G. Colombo 3, Padova, 35121, ITALY

FEATURES

source

1..1756

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="1"

/map="p35.1-35.2"

/clone="STRAIT03036"

/tissue\_type="skeletal muscle"

/clone\_lib="pEAK8"

/dev\_stage="adult"

1..1756

/gene="MURF2"

1..282

/gene="MURF2"

/number=1

124..1185

/gene="MURF2"

/codon\_start=1

/product="muscle specific RING finger 2 (MURF2)"

/protein\_id="CAC81706.1"

/db\_xref="GI:18073356"

/translatio="ADYSSSSLODNGMSPRCEHEDICPLEKNEFTRVALLPOCHNI  
 RKANDIHOANPWNTWYSSSGRGRPCREHVIDNADVYGLONLIVENT  
 DIYOGCSPLONGSHPCREHDEPITINUYICTCVPSCMGKVEIRHACFAPIA  
 SVFOSQRTLENNICMLVANDRYOTITIOLESRYTRKNSHVOYELSOXPTIKAX  
 LIDRKSBLRTOIOEQEKRSITELIOTIOLOLKRKLTVETALVIOLESBGATIL  
 LTAQALIKSIVAYSAGCQLOEGEENDFILIDELIADALRAIDFGDDEEEFIF  
 DEEDOEFEESSTEGKREBGHQ"

BASE COUNT	466 a	398 c	516 g	376 t
exon	283. .455	/gene="MURF2"		
	/number=2			
exon	456. .624	/gene="MURF2"		
	/number=3			
exon	625. .720	/gene="MURF2"		
	/number=4			
exon	725. .954	/gene="MURF2"		
	/number=5			
exon	955. .977	/gene="MURF2"		
	/number=6			
exon	978. .1102	/gene="MURF2"		
	/number=7			
exon	1103. .1174	/gene="MURF2"		
	/number=8			
exon	1175. .1756	/gene="MURF2"		
	/number=9			
polyA_signal	1720. .1725	/gene="MURF2"		
polyA_site	1756	/gene="MURF2"		





Db	318	TCGATGCTGGAGGCGCTTTCGGCTGGCCCCACCTGGCCGCGCCACGAGGTGATCATGATGATCGT	377
Qy	90	HISGLYVALTYRGLYLEUGINARGASPLEUENLEUVALGLUASNILLEASPILLETYRGLYS	109
Db	378	CACGAGCTGTAAGGCGCTGCAGAGAACCTGCTGGTGGGAACATCATCTGACATCTACACAA	437
Qy	110	GLINGLU---SERSERATGPROLEUHNISALALYSALAGLUGLINFHNISLEUMETCYSGLU	128
Db	438	CAGAGTGTCTCAACTGCGCGCTG-----CAGAAAGGCACTCAACCCATGTGCAAGAG	491
Qy	129	HISGLUASPGULYSILLEANILLETYRCYSGLEUSERCYSGULVALPROTHCYSERLEU	148
Db	492	CACCAAAATGGAAATCAACATCTACTGTCATCGTGTAGGTGAGTCCACCTGCTCCATG	551
Qy	149	CYSRYSVALPHEGLYALAHISLYSASPQSYGLUVALALAPROLEUPROTHILETYRGLYS	168
Db	552	TGCAAGGTGTTGGGATTCACAAAGCCCTGGAGGTTGGGCCCATTCGACAGTGTCTTCAG	611
Qy	169	ARGGLNYSERGLULEUSERASPGLYILLEALMETLEUVALALAGLYASNASPARVAL	188
Db	612	GCACAAAAGACTGGAACGAATAATCATGTATCTCCATGCTGCTGGGGGAATGACCGTGTG	671
Qy	189	GLNLAVALIIEHNGIMETGSLUGLNUVALCYSGINTHRIIEGLIUNASPARSERARG	208
Db	672	CAGACCATCATCTACCTGAGCTGGAGGATTCGCCGTGAGTGACCAAGGAADAAGTACACAG	731
Qy	209	GLNYSGLINLEULEUASNGLINARGPHEGLIUNHIREUCYSALAVALLLEUGLUNARGLYS	228
Db	732	GTAAGGAGAGAGCTGAGCCAGAGATTGTACACGTTGTATGTCATCTCGATGAGAGAA	791
Qy	229	GLYULULEUENGLINLALALEUALARGGLUGLUGLUGLULYSLEUGLINARGVALARGELY	248
Db	792	AGTGAAGTTCGCGACCGGATCACCGAGGAGGAGGAGGAAAGCTTAATCATCCAGGCC	851
Qy	249	LEULIARGGLINTYRGLYASPHISLEUGLUNGLYSERSERLYSEUVALGLUSERALALE	268
Db	852	CTCATCCAGCAGATGCCAGAGCAGCAGCTGGACAAAGTCCAAAGAGCTGTGTGAATCTCCATC	911
Qy	269	GLNSERNETGLUUPROGLIMETALALEUTYRLEUGINGLINALALYSGLULEULEASN	288
Db	912	CAGTCCCTGGGACGAGCCCTGGGGAGCCACTTCTCTTCATGCTGCACAACTCATCTCAA	971
Qy	289	LYSVALGLYALAMETSERLYSVALGLULEUENLAGLYARGPROGLINTYRGLIUSER	308
Db	972	AGCATTTGGGAAGCTTCCAAAGGCTGCCACGCTGGGGAAAGACAGACAGAGGCTTTAGAAC	1031
Qy	309	METGLUINPHESERVALSERVALGLUHNISVALALAGLUMETLEUARGTYHRIEASPH	328
Db	1032	ATGGACTCTTACTTGTGATTAGAGCACATAGACAGCGCCCTGAGAGCCATTGACTTT	1091
Qy	329	GLNPROGLYALAHLAGLYASPGULUNSPASPMETALA-----LEUASP	344
Db	1092	-----GGGACGAGTAGGAGAAAGGAAAGAAATTCATTGAAGAAAGATCAGAA	1139
Qy	345	GLYGLUGLUGLYASNALAGLYLEUGLUNLU	354
Db	1140	GAGAGAGAGTCCACAGAGAGGAGAGANADA	1169
RESULT 20			
LOCUS	AB047601		
DEFINITION	Macaca fascicularis brain cDNA, clone: Qnpa-10466.		
ACCESSION	AB047601		
VERSION	AB047601.1		
KEYWORDS	GI:9929936		
SOURCE	Macaca fascicularis		
ORGANISM	Macaca fascicularis		
REFERENCE	1		
AUTHORS	Osada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirata, M.,		

[illegible]





DEFINITION BB140247 RIKEN full-length enriched, adult male bone Mus musculus  
 cDNA clone 9830168H24 3', mRNA sequence.  
 ACCESSION BB140247  
 VERSION BB140247.1 GI:8795184  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 316)  
 AUTHORS Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci  
 P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,  
 Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,  
 Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,  
 Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,  
 Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
 Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata  
 Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y.,  
 Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tominaga, N., Toya  
 T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I.,  
 Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino  
 M., Muramatsu, M. and Hayashizaki, Y.  
 TITLE RIKEN Mouse ESTs (Konno, H., et al.)  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 FEATURES  
 source Location/Qualifiers  
 1. 316  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="9830168H24"  
 /clone\_lib="RIKEN full-length enriched, adult male bone"  
 /sex="male"  
 /tissue\_type="bone"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Site\_1: SalI; Site\_2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTT 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. cDNA went through one round of normalization  
 to Rot = 10.0 and subtraction to Rot = 185.0. Second  
 strand cDNA was prepared with the primer adapter of  
 sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC  
 3']. cDNA was cleaved with XhoI and BamHI. Vector: a  
 modified pBluescript KS(+) after bulk excision from Lambda  
 FLC I."  
 BASE COUNT 67 a 91 c 105 g 53 t  
 ORIGIN  
 Query Match 7.5%; Score 108; DB 10; Length 316;  
 Best Local Similarity 99.4%; Pred. No. 1.3e-40;  
 Matches 158; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1198 GGGGATGAAGAGGATGACGACATGGCTTTGGATGGGGAGGAGGGCAATGCGGGGCTGGAG 1257  
 Db 81 |||||  
 QY 1258 GAGGAGCGGCTGGACGTGCCAGAAGGCTCAGGCCTGCACTGACCCGACTCTGATCCAGAG 1317  
 Db 141 |||||  
 QY 1318 CGCACACCCGAAGCGGGAGCCAAGGGATGCTGAGGATCT 1356  
 Db 201 |||||  
 CGCACACCCGAAGCGGGAGCCAAGGGATGCTGAGGATCT 239

